REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree PhD Year 2005 Name of Author ONO, S.

The Biophysical Characterisation of the Enterobacterial Nucleoid Proteins H-NS and StpA

Shusuke Ono

Department of Biochemistry and Molecular Biology
University College London
Gower Street
London

For the degree of Doctor of Philosophy (PhD)

May 2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

To my Mother and Father

Abstract

H-NS is a major protein component of the nucleoid, found in many Gram-negative bacterial species. H-NS is involved in the modulation of expression of a wide range of genes, as well as contributing towards nucleoid structure. There are two structurally independent domains in H-NS, one involved in protein-protein interactions (the N-terminal oligomerisation domain), and the other involved in DNA-binding (C-terminal). These are joined via a flexible linker sequence. H-NS both self-associates and interacts with other nucleoid-associated proteins to form specific oligomeric complexes that bind DNA, allowing a precise level of control in gene expression.

The protein StpA, a paralogue of H-NS with 58% sequence identity, was originally identified by its RNA chaperone activity. Subsequent studies have suggested structural similarities between H-NS and StpA, with a degree of overlap in their function *in vivo*. StpA self-associates in a similar manner to H-NS, and has been shown to interact with H-NS. Whilst small differences in the properties of H-NS and StpA have been identified, no clear distinctions have been made between the two proteins.

This work investigates several key issues regarding the properties of the StpA protein. A number of biophysical techniques have been used to investigate the interaction of H-NS and StpA. The results of these experiments are consistent with a model whereby StpA self-associates via a 'head-to-tail' interaction. Furthermore, StpA exhibits a different affinity and kinetic behaviour of association in comparison to H-NS. The properties of self-association and interaction of H-NS and StpA are fully consistent with studies that highlight the intimate relationship between the two proteins *in vivo*.

Two independent structures of the N-terminal oligomerisation domain of H-NS have been reported. These structures were derived at different temperatures (i.e. 25 °C and 35 °C). To investigate the implications of these model structures on the thermoregulatory functions of H-NS, the oligomerisation properties of H-NS were investigated over a temperature range. Both the oligomerisation and DNA-binding properties of H-NS were found to vary within a physiologically relevant range of temperatures.

To characterise the interaction of StpA with DNA and to allow comparison with H-NS, the solution structure of the C-terminal domain of StpA was determined, solved by NMR. The interaction of DNA with this domain was characterised using both NMR and calorimetric methods.

Statement

The work described in this thesis was carried out in the Department of Biochemistry and Molecular Biology, University College London between 2000 and 2004. The NMR spectra were acquired with the assistance of Dr. M. Williams, Dr. R. Harris or J. Taylor. All experiments involving H-NS_{1-89} were conducted with the assistance of T. Olsson. All other work was carried out by the author. This project was funded by the Biotechnology and Biological Sciences Research Council (BBSRC).

Some of the work detailed in this thesis has been published elsewhere:

Esposito, D., Petrovic, A., Harris, R., Ono, S., Eccleston, J. F., Mbabaali, A., Haq, I., Higgins, C. F., Hinton, J. C., Driscoll, P. C., and Ladbury, J. E. (2002) H-NS oligomerisation domain structure reveals the mechanism for high order self-association of the intact protein. *J Mol. Biol.* **324**, 841-850.

Acknowledgements

I would like to thank all of my colleagues and friends, whose kind assistance and expertise I have relied upon throughout my years at UCL. I owe a debt of gratitude to Dr. Mark Williams, without whom none of the NMR studies would have been possible. In particular I would like to thank Dr. Matthew Cliff and Jonathan Taylor who patiently provided me with the expertise and confidence to deal with many aspects of the project both inside and outside the laboratory. I would also like to mention Dr. Simon Bergqvist, Radwan Fawaz, Brian Ferguson, Acely Garcia, Dr. Roger George, Tjelvar Olsson, Dr. Arsen Petrovic and Dr. David Vines, who helped me at almost every stage of my project and provided invaluable insights into the many problems that I faced. Acknowledgement is also due to my academic mentor, Dr. Paul Driscoll. I also wish the greatest possible successes to Paul Leonard with his future work with H-NS and StpA.

I would like to extend my appreciation to everyone with whom I have shared my exciting adventure through the 'PhD labyrinth'. I have made many priceless friendships and happy memories for which I am grateful. Rachael Lindup in particular has helped me retain my sanity by her resolute support and sympathy. I would also like to thank my wonderful family for their encouragement and unwavering support.

Last but not least, I would like to sincerely thank Prof. John Ladbury for his supervision and guidance throughout this project, for the respect with which he treated me, for the confidence which he instilled in me and for being a damn good bloke.

Table of Contents

Chapter 2 – Materials and methods

List of figures

List of tables

Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
Å	Angstrom
A	Adenine
ARIA	Ambiguous Restraints for Interactive Assignments program
AUC	Analytical ultracentrifugation
bp	base pairs
°C	Degrees Celsius
C	Cytosine
C-terminal	Carboxyl-terminal
CD	Circular dichroism spectroscopy
ΔC_p	Change in heat capacity
ΔH_{obs}	Observed enthalpy of binding
DNA	Deoxyribonucleic acid
FID	Free induction decay
G	Guanine
H-NS	histone-like nucleoid structuring protein
$H\text{-}NS_{1\text{-}64}$	H-NS truncation mutant with first 64 residues
$H\text{-}NS_{1\text{-}89}$	H-NS truncation mutant with first 89 residues
$H\text{-}NS_{FL}$	Full-length H-NS
HSQC	Heteronuclear single quantum coherence
IPTG	isopropyl-B-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_b	Association constant
K_d	Dissocitaion constant
kDa	Kilo Dalton
LB	Luria-Bertani

mRNA	Messenger RNA
MW	Molecular weight
n	Stoichiometry of binding
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
N-terminal	Amino-terminal
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point
RF	Radio frequency
RMSD	Root mean square deviation
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
StpA	suppressor of the *td* mutant phenotype A
$StpA_{1-65}$	StpA truncation mutant with first 65 residues
$StpA_{91-134}$	The C-terminal domain of StpA
$StpA_{FL}$	Full-length StpA
T	Thymine
T_1	Longitudinal relaxtion time constant
T_2	Transverse relaxation time constant
T_m	Melting temperature
TOCSY	Total correlation spectroscopy
UV	Ultraviolet

Chapter 1

Introduction

1.1 The bacterial nucleoid

Unlike eukaryotic cells, bacteria do not have a nucleus nor defined structures such as nucleosomes (Ussery *et al.*, 1994); nevertheless the bacterial genome is condensed into a dynamic structure known as the nucleoid (Robinow, 1956; Sherratt, 2003). The nucleoid allows the packing of chromatin into a fraction of the cellular volume, and varies in size and shape according to growth conditions (Hirschbein and Guillen, 1982). The dynamic nature of the nucleoid is reflected in the large variation of abundance of the nucleoid-associated proteins during different stages of the *Escherichia coli* growth cycle (Ali Azam *et al.*, 1999).

The *E. coli* nucleoid is associated with a relatively small number of highly abundant DNA-binding proteins, such as CbpA (curved DNA-binding protein A), Dps (DNA-binding protein from starved cells), H-NS (histone-like nucleoid structuring protein), HU (heat unstable protein), IHF (integration host factor), Fis (factor for inversion stimulation), Hfq (host factor for Q_β phage) and StpA (suppressor of the *td* mutant phenotype A). These proteins are present in the cell in their tens of thousands of molecules per cell (Ali Azam *et al.*, 1999). However, none of these nucleoid-associated proteins bear any sequence or structural homology with the eukaryotic his-

tones, and the mechanisms of bacterial DNA-packaging and organisation are different from those of eukaryotes (Rimsky *et al.*, 2001). Indirect immunofluorescence spectroscopy has shown the major nucleoid-associated proteins H-NS, HU, IHF, StpA and Dps are uniformly distributed throughout the nucleoid, consistent with their main role as proteins that provide structure and organisation to the nucleoid (Azam *et al.*, 2000). These proteins have been referred to as histone-like, due to the similarity of nucleoid-associated proteins to histones in properties such as DNA-binding, high abundance, basic charge and low molecular mass; however there is no resemblance in amino acid sequence or structure (Dorman and Deighan, 2003; Dorman, 2004).

Recent atomic force microscopy (AFM) studies suggest the shape and size of the nucleoid is determined by the interplay of these nucleoid-associated proteins. For instance, H-NS has been shown to promote global compaction of DNA, whereas HU antagonises compaction by 'opening up' DNA structure. The ratios of these two proteins have been suggested to contribute towards the overall compactness of the nucleoid (Dame *et al.*, 2000; Dame and Goosen, 2002). However, the precise mechanisms of nucleoid compaction and organisation by these nucleoid-associated proteins is unknown (Sherratt, 2003; Dorman and Deighan, 2003). The details of the two proteins of interest to this thesis, H-NS and StpA, will be introduced in the remainder of this chapter.

1.2 Overview of the histone-like nucleoid structuring protein (H-NS)

The histone-like nucleoid structuring protein (H-NS) has been identified as a major component of the *E. coli* nucleoid, with its abundance measured in the tens of thousands per cell (Free and Dorman, 1995; Ali Azam *et al.*, 1999). The term 'histone-like' was used loosely to describe H-NS properties and function; however there is no sequence or structural similarity between H-NS and histones (Dorman and Deighan, 2003; Dorman, 2004). H-NS and related proteins appears widely distributed amongst the α, β and γ subdivisions of proteobacteria (Tendeng and Bertin, 2003). A paralogous protein, StpA, has a high level of sequence identity (58%) and displays many similarities with H-NS structure and function (Zhang and Belfort, 1992).

Consistent with its role as a nucleoid-structuring protein, H-NS has been shown to affect DNA topology and constrain negative supercoils (Owen-Hughes *et al.*, 1992;

Tupper *et al.*, 1994). Overexpression of H-NS led to global repression of transcription and caused an extended period of arrested growth (McGovern *et al.*, 1994). In a separate study, overexpression of H-NS led to a strong and immediate inhibition of transcription and translation, and to a lesser degree, inhibition of DNA and cell wall synthesis. Cell viability was drastically reduced, possibly because of the above effects. Immunoelectron microscopy indicated that the nucleoid was considerably compacted into dense and almost spherical 'nuclei' as a result of the overexpression of H-NS (Spurio *et al.*, 1992). AFM of H-NS/DNA complexes suggest that H-NS compacts plasmid DNA into dense globular structures by lateral condensation of DNA (Dame *et al.*, 2000; Dame *et al.*, 2001). The biological significance of the hairpin-like structures of H-NS/DNA complexes caused by lateral condensation of DNA has been demonstrated at the H-NS-dependent repression of the *rrnB* P1 promoter (Dame *et al.*, 2002).

In addition to the role of structuring the nucleoid, H-NS functions as a global regulator of transcription, conferring the cell with metabolic diversity and the ability to adapt to a wide range of environments (Atlung and Ingmer, 1997; Tendeng and Bertin, 2003). H-NS was identified by the high pleiotropism of *hns* mutant phenotypes (Ussery *et al.*, 1994; Atlung and Ingmer, 1997; Ussery *et al.*, 2001). Gene expression profiling using DNA arrays have revealed that there may be as many as 250 genes under regulation by H-NS. The expression over 80% of these genes were suggested to be negatively modulated by H-NS, supporting the current view that H-NS is a global repressor of gene expression (Hommais *et al.*, 2001).

The *hns* gene, encoding the H-NS protein, has been mapped to the 27.5 min region on the *E. coli* chromosome (Goransson *et al.*, 1990). H-NS is a heat-stable, 15.6 kiloDalton (kDa) polypeptide with 137 amino acid residues. H-NS is neutral, with an isoelectric point (pI) of 7.5 (Ussery *et al.*, 1994). It binds preferentially to curved DNA, compared to moderately or non-curved DNA, independent of DNA sequence (Zuber *et al.*, 1994; Jordi *et al.*, 1997; Azam and Ishihama, 1999; Poore and Mobley, 2003). It also forms higher order oligomers, a property crucial for its function as a transcriptional repressor (Williams *et al.*, 1996; Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997; Smyth *et al.*, 2000; Badaut *et al.*, 2002).

1.3 Phylogenetic distribution of H-NS

More than seventy H-NS and related proteins have been found to date, widely distributed amongst Gram-negative bacteria (Bertin *et al.*, 2001; Tendeng and Bertin, 2003). The distribution of H-NS and related proteins appears to be limited to the α, β and γ subdivisions of proteobacteria, according to the 16S rRNA (ribosomal ribonucleic acid) classification system (Tendeng and Bertin, 2003). H-NS from species from the γ-proteobacteria subdivision, order enterobacteriales, family enterobacteriaceae, such as *E. coli* and *S. typhimurium*, and StpA from *E. coli*, represent the most studied members of the H-NS-like proteins to date.

Studies of the structural and functional domain organisation of the H-NS-like proteins genes suggest this family includes SPB (a transcription factor) from *Rhodobacter sphaeroides*, XrvA (regulates virulence to rice) from *Xanthomonas oryzae*, and VicH from *Vibrio cholerae*, although there are differences in their physiological roles and genetic organisation (Shimada *et al.*, 1996; Bertin *et al.*, 1999; Tendeng *et al.*, 2000). Other examples of a *hns*-like genes include BpH3 from the non-enteric *Bordetella pertussis*, the aetiological agent of whooping cough, and HvrA from the non-pathogenic *Rhodobacter capsulatus* (Buggy *et al.*, 1994; Goyard and Bertin, 1997). H-NS-like proteins have also recently been identified in two cold-adapted bacteria *Acinetobacter* and *Psychrobacter* (isolated from Siberia and Antarctica respectively). The H-NS-like protein from the *Psychrobacter* displayed significant differences in thermal stability compared to other members of the H-NS-like family of proteins, suggesting a structural adaptation of the protein to lower temperature (Tendeng *et al.*, 2003).

Sequence comparisons of all known *hns*-like genes to date suggest the carboxyl-terminal (C-terminal) domain is well conserved, whereas the amino-terminal (N-terminal) domain is less so. The linker region connecting the two domains is not conserved (Tendeng and Bertin, 2003). The H-NS-like family includes proteins which only possess the N-terminal or C-terminal domain of H-NS, highlighting the modular nature of H-NS. Studies on the *hly* operon have revealed the sequence similarity of *E. coli* Hha (and the paralogous YmoA from *Yersinia enterocolitica*) to the N-terminal domain of H-NS, suggesting that Hha and YmoA are distant paralogues of H-NS, however, these proteins lack an equivalent to the C-terminal domain of H-NS (Nieto

et al., 2002). However comparison of the three dimensional structures of Hha and the two structures of N-terminal of H-NS, all determined by nuclear magnetic resonance (NMR) spectroscopy, reveals significant differences in structure between Hha and H-NS, casting doubt on the suggestion by Nieto *et al.* that H-NS and Hha are related (Esposito *et al.*, 2002; Yee *et al.*, 2002; Bloch *et al.*, 2003). The N- and C- terminal halves of the DNA-binding protein KorB are very similar to each other, suggesting *korB* arose by gene duplication. Interestingly, both halves showed significant homology to the C-terminal DNA-binding domain of H-NS and the related protein HvrA from *R. capsulatus* (More *et al.*, 1996). The protein Ler (LEE-encoded regulator) is a transcriptional activator of the H-NS-repressed LEE pathogenicity island of enteropathogenic *E. coli* (EPEC) (Umanski *et al.*, 2002; Haack *et al.*, 2003). Ler is a 15 kDa protein that displays homology to H-NS, with more sequence similarity to H-NS in the C-terminal domain than the N-terminal domain (see sequence alignment in Section 1.9) (Bertin *et al.*, 2001; Haack *et al.*, 2003). Like H-NS, mutations that disrupt the formation of the coiled-coil in the N-terminal region of Ler were shown to disrupt Ler function and DNA-binding (Sperandio *et al.*, 2000). Interestingly, whilst Ler displays characteristics common with H-NS, such as the ability to oligomerise and to bind DNA sequences with intrinsic curvature, Ler was shown to be a specific transcriptional activator of virulence genes, by displacing H-NS and antagonising the repressive effect of H-NS (Bustamante *et al.*, 2001; Haack *et al.*, 2003). The antagonistic effect of Ler was not observed at other H-NS-dependent operons such as *bgl* (Haack *et al.*, 2003). The DNA-binding protein MdbA from uropathogenic *E. coli* (UPEC) was originally identified as a member of the H-NS family by its high degree of homology with the N-terminal domain of H-NS, with a C-terminal domain that was not related (Cusick and Belfort, 1998). However subsequent analysis revealed that a single base pair insertion at nucleotide 405 produces an open reading frame with full-length homology to H-NS (Dorman *et al.*, 1999).

It is worth noting that approximately half of the relatively large genome of *E. coli* contains paralogous genes (when analysing 75% of the genome). In contrast, 35% of the genes from *Haemophilius influenzae* and 17% of genes from *Helicobacter pylori* contain paralogous genes. It has been suggested that the high proportion of paralogous genes in *E. coli* reflects its ability to adapt to a wide range environments (Hinton, 1997).

1.4 Target genes for H-NS control

The H-NS-like family of proteins have been attributed with a variety of vital roles in the bacterial cell, and mounting evidence points to the universal role of these proteins in conferring the cell with metabolic diversity and the ability to adapt to a wide range of environments (Atlung and Ingmer, 1997; Tendeng and Bertin, 2003). Consistent with its role in the cell as a global regulator of gene expression, H-NS was first identified by the highly pleiotropic nature of *hns* mutant phenotypes (Ussery *et al.*, 1994; Atlung and Ingmer, 1997; Ussery *et al.*, 2001).

Whole proteome and transcriptome analyses have revealed that the expression of approximately 5% of genes (corresponding to 250 genes) was altered, directly or indirectly, in an *hns* mutant. The view of H-NS as a repressor of gene expression is reinforced by the observation that the expression of over 80% of these genes was increased as a result of the *hns* mutation. Many of these H-NS-dependent genes are involved in mediating responses to changes in environment, for instance changes in oxygen availability, temperature, osmolarity, and changes in pH. Figure 1.1 shows a classification of the H-NS-dependent genes. The authors concluded H-NS played a vital role in maintaining intracellular homeostasis (Hommais *et al.*, 2001). Table 1.1 shows a list of genes whose regulation has been shown to be modulated by H-NS. Whilst cases have been identified where the expression of a gene was found to be under positive transcriptional regulation by the involvement of H-NS, for example expression of *traJ* and *malT* (Johansson *et al.*, 1998; Starcic-Erjavec *et al.*, 2003), there is no evidence for H-NS acting directly as a transcriptional activator (Dorman, 2004).

The role of H-NS in the modulation of expression of genes involved in adaptation to environmental parameters is well documented. For instance, in *Salmonella enterica* serovar Typhimurium, H-NS is involved in the acid tolerance response, by repressing the transcription of the *ompR* acid-induced response regulator. In an *hns* mutant, the derepression of *ompR* results in a constitutive acid tolerance response (Bang *et al.*, 2002). Compared to wild-type, *hns*-mutants of *E. coli* are more resistant to low pH (Bertin *et al.*, 2001). By modulating the expression of flagellar genes and liposaccharide production, H-NS has been found to regulate the adhesion of *E. coli* to solid hydrophilic surfaces, in response to anaerobic conditions. Regulation of several porin genes by H-NS was also observed (Landini and Zehnder, 2002). H-NS and StpA

were shown to thermoregulate the maltose regulon by stimulating translation of maltose regulon activator MalT (Johansson *et al.*, 1998).

Many of the H-NS-like family of proteins have been implicated in the modulation of the expression of genes involved in bacterial virulence. The regulation of bacterial virulence is intimately linked with the cell's ability to adapt to changes in environment as the conditions inside a host can differ significantly in osmolarity, pH, temperature, and oxygen availability, compared to other ecological circumstances. Interestingly, the majority of the members of the H-NS-like proteins identified to date are derived from pathogenic species. Furthermore, species that have more than one *hns*-like gene show a wide ecological range and metabolic diversity, suggesting H-NS-like proteins are involved in adaptation to a diverse range of environments (Tendeng and Bertin, 2003). For example, enteropathogenic *E. coli* possess a unique 35.6 kbp pathogenicity island called the locus of enterocyte effacement (LEE). The LEE is associated with the formation of attaching and effacing (AE) lesions in epithelial cells, which leads to severe diarrhoea in children. The LEE virulence genes are thermoregulated by H-NS, being repressed at 27°C and expressed at 37°C (Umanski *et al.*, 2002). H-NS has been shown to mediate the gene expression of the *flhDC* master operon, which governs motility, chemotaxis, and differentiation of enterobacteria into swarming cells (Soutourina *et al.*, 2002). In *Erwinia chrysantemi*, *pecT* encodes a regulatory protein that represses the expression of pectate lyases (Pel) that degrade plant cell walls. H-NS itself exerts repression of gene expression on *pecT*: the H-NS/Pel system presents a rare case where H-NS has a positive effect (via the second regulator PecT) on gene expression (Nasser and Reverchon, 2002). The involvement of H-NS was demonstrated in the thermoregulation of the urease gene cluster that encodes the *Proteus mirabilis* urease virulence factor, leading to differential expression of urease at 25° C and 37 °C (Poore and Mobley, 2003).

Gene	Function
appY	Anaerobic growth phase activator
bgl	B-glucosidase
bolA	Cell shape regulator
cfaD	Pili formation
chiA	Endochitinase
clyA	Cytolysin A
csgA	Curli formation
csiD	C-starvation induced genes
cspA	Major cold-shock protein
cydAB	Cytochrome d oxidase
espADB	Enterohemorrhagic *E. coli*
fimA	Type 1 fimbrial subunit
fimB	Recombinase for *fimA* promoter flipping
flhD	Flagellar control protein
fliA	Flagella-specific σ factor
gadABC	Glutamic acid decarboxylase
hdfR	Transcriptional regulator of the flagellar master operon *flhDC*
hlyCABD	Haemolysin
hns	DNA structuring proiein, transcription factor
kpsF	Capsule gene
lrp	Leucine-responsive regualtory protein
luxC	*luxC, luxRI* region of luminous bacteria
lysU	Lysyl-tRNA-synthetase
malEFG	Maltose regulon
malT	Activator of maltose regulon
micF	Antisense RNA
nirB	Nitrate reductase
ompC	Outer membrane protein
pap	Pili formation
proU	Proline/glycine betaine transport system
rcsA	Activator of capsular polysaccharide synthesis
rpoS	Stationary σ factor
rrn	Ribosomal RNA
stpA	Possible molecular backup for H-NS, RNA chaperone activity
toxR	Virulence gene regulatory protein (*V. cholerae*)
ureR	AraC-like activator (*P. mirabilis*)
virB	Invasion regulatory gene (*S. flexneri*)
yebG	Member of the SOS regulon, unknown function

Table 1.1
A list of genes whose expression is known to be modulated by H-NS. This list was adapted from Schröder and Wagner, 2002.

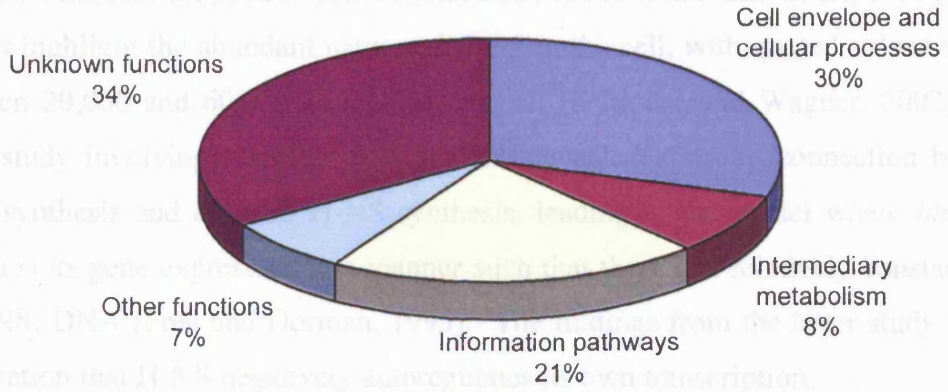


Figure 1.1
Classification of the (approximately 250) genes under H-NS regulation. Genes that displayed significant differential expression between wild-type and *hns* mutant strains of *E.coli*, determined by expression profiling using DNA arrays, were classified according to their function. This figure was adapted from Hommais *et al.* 2001.

1.5 H-NS fulfils a variety of roles in the cell

In addition to the role as a modulator of gene expression, H-NS has been attributed with other functions. For example, H-NS and Lrp have been shown to activate *traJ* expression in *E. coli*. TraJ is the main positive regulator of the *tra* operon, which encodes the function of conjugative transfer of F-like plasmids (Starcic-Erjavec *et al.*, 2003). H-NS was found to play a role in the modulation of the site-specific inversion rate of the 314 bp *pilA* promoter (also known as *fimA*) of the gene that encodes the type 1 pilin monomer (Kawula and Orndorff, 1991; O'Gara and Dorman, 2000). IS*1* encodes a transposase, and is also an active transposable element present in genomes and plasmids of enteric bacteria. Whilst H-NS did not affect transcription from the promoter of IS*1*, it was found to be important in the transpositional recombination events at the IS*1* ends (Shiga *et al.*, 2001). H-NS has been shown to exhibit RNA chaperone activity *in vitro*, although the *in vivo* significance of this observation is not clear (Mayer *et al.*, 2002).

1.6 Regulation of H-NS

H-NS negatively autoregulates its own transcription (Ueguchi and Mizuno, 1993; Ueguchi *et al.*, 1993; Falconi *et al.*, 1993; Dersch *et al.*, 1993; Free and Dorman, 1995; Badaut *et al.*, 2002). In one study involving quantitative Western blot analysis, H-NS was found to be maximally expressed during the exponential growth phase,

with as many as 20,000 molecules per cell. On entering the late stationary growth phase, the number dropped to 8,000 molecules per cell (Ali Azam *et al.*, 1999). Other studies highlight the abundant nature of H-NS in the cell, with quoted values varying between 20,000 and 60,000 molecules per cell (Schroder and Wagner, 2002). Another study involving Northern blot analysis revealed a strong connection between DNA synthesis and *de novo* H-NS synthesis, leading to the model where *hns* autoregulates its gene expression in a manner such that there is a relatively constant ratio of H-NS: DNA (Free and Dorman, 1995). The findings from the latter study fits the observation that H-NS negatively autoregulates its own transcription.

Changes in environmental parameters do not usually result in a change in *hns* expression (Soutourina *et al.*, 2002). The upregulation of *hns* transcription in response to cold-shock is an exception. The CspA protein is involved in the transcriptional activation of some proteins of the cold-shock response, and mediates the 3- to 4-fold enhancement of transcription of *hns* in response to a cold shock. *E. coli* carrying *hns* mutations have been shown to have a severely impaired ability to adapt to a cold environment (Dersch *et al.*, 1994). It has been suggested that the upregulation of H-NS in the cold-shock response merely ensures intracellular homeostasis, rather than to affect global transcription. Lower temperatures would thermodynamically favour smaller oligomeric forms, and an increased intracellular concentration of H-NS would maintain the optimum oligomeric state needed for full H-NS function at lower temperatures (Ceschini *et al.*, 2000).

Several H-NS binding sites have been identified upstream of the *hns* promoter. Upon binding of H-NS to the two high-affinity sites, the lower affinity site is cooperatively filled by H-NS. This lower affinity site partially overlaps the promoter, occluding RNA polymerase access to the -35 box, thus effecting repression (Falconi *et al.*, 1993). These observations agree closely with the 'nucleation' model (discussed in Section 1.9.6) of H-NS-mediated repression of gene expression (Rimsky *et al.*, 2001; Badaut *et al.*, 2002). The major nucleoid proteins FIS and H-NS were shown to be antagonistic at the *hns* promoter, possibly by competing for the same binding sites. At least seven FIS binding sites were discovered in the *hns* promoter region, partially overlapping high-affinity H-NS binding sites involved in *hns* autoregulation. The growth-phase dependence of the intracellular concentrations of H-NS and FIS was

then suggested to contribute towards the modulation of *hns* expression (Falconi *et al.*, 1996). The genes *hns* and *stpA* were shown to exert parallel autogenous control and cross-regulation, with the overexpression of StpA leading to repression of transcription from *hns* (Zhang *et al.*, 1996; Sonden and Uhlin, 1996).

A small, stable untranslated RNA molecule DsrA, consisting of 87 nucleotides, has been implicated in the post-transcriptional regulation of H-NS. DsrA is thought to have a global regulatory role by affecting the transcription of H-NS and the stationary-phase and stress response sigma factor RpoS (also known as σ^S). H-NS and RpoS are apparent antagonists of the modulation of expression of a variety of stress response genes. DsrA binds both the 5' and 3' regions of *hns* mRNA, and inhibits *hns* translation by blocking initiation of translation and by enhancing turnover of the *hns* mRNA. The translation of *rpoS* mRNA is enhanced when DsrA binds a *cis*-acting inhibitory loop of *rpoS* mRNA, thus exposing the initiation sequence of *rpoS* mRNA. The regulation of *dsrA* itself is subject to a number of proteins, including H-NS and StpA (Lease *et al.*, 1998; Lease and Belfort, 2000).

1.7 Interaction with other proteins

The mechanism of repression of H-NS varies according to target genes, with some promoters requiring proteins other than H-NS for full regulatory control. It has been suggested that heterooligomeric complexes with varying proportions of nucleoid-associated proteins have different regulatory properties (Dorman *et al.*, 1999). Indeed, the identification of many proteins shown to interact with H-NS suggests that the global regulation of gene expression by H-NS requires the formation of specific nucleoprotein complexes depending on the promoter.

The interaction of H-NS with the paralogous protein StpA (with 58% sequence identity) is well documented. Dominant-negative mutants of H-NS have been shown to disrupt the activity of wild-type StpA, and vice-versa, suggesting H-NS and StpA form heteromeric complexes. This implication was supported by the observation of cross-linked species of H-NS and StpA heteromers (Williams *et al.*, 1996). Studies of Lon-protease-resistant mutations of StpA suggest StpA is present in *E. coli* in a heteromeric form with H-NS (Johansson and Uhlin, 1999).

Interestingly, three highly homologous proteins from the H-NS family have been identified, H-NS, StpA and Sfh, in *Shigella flexneri* 2a strain 2457T. All three proteins were shown to auto-repress themselves and cross-repress the expression of the other proteins. Sfh was shown to form heterodimers with both H-NS and StpA. A hypothesis was presented, where heterooligomers with different combinations of H-NS, StpA and Sfh display different regulatory properties, allowing a high level of control over cellular processes by expressing varying ratios of the proteins (Deighan *et al.*, 2003; Beloin *et al.*, 2003). *Pseudomonas putida* has recently been identified as having 5 *hns*-like genes, suggesting global regulatory networks under the influence of several *hns*-like genes are common place (Tendeng and Bertin, 2003). As paralogues from the H-NS gene family are often acquired via genetic exchange, for example via plasmids and pathogenicity islands, it has proposed that these paralogues have a regulatory role over the genetic elements that encode them (Deighan *et al.*, 2003).

Protein complexes formed by the interaction of H-NS and its paralogues have been shown to specifically regulate gene expression. For example, the *E. coli* hemolysin operon (*hly*) is normally repressed under conditions of high osmolarity and at low temperatures by a complex of the proteins Hha and H-NS (Nieto *et al.*, 2002). Reminiscent of the *virF* operon (Falconi *et al.*, 1998a), there are two H-NS binding sites upstream of the *hly* promoter. Hha does not show any specificity for DNA sequence. Loss of either Hha or H-NS results in the derepression of *hly* (Nieto *et al.*, 2000; Madrid *et al.*, 2002).

The fortuitous copurification of H-NS with T7 gene *5.5* protein led to studies that showed the T7 gene *5.5* suppresses the repressive effect of H-NS on transcription by binding specifically and tightly to H-NS *in vitro* and *in vivo*, suggesting a role for this interaction in aiding reproduction of T7 bacteriophage in *E. coli* (Liu and Richardson, 1993).

1.8 Physiochemical aspects of H-NS

E. coli H-NS is a 15.6 kDa protein with 137 amino acid residues and is neutral under native conditions, with a pI of 7.5 (Ussery *et al.*, 1994). There are three distinct isoforms of *E. coli* H-NS that exhibit different isoelectric points, although the biological significance of this observation is not yet clear (Ussery *et al.*, 1994). Post-translational modification of H-NS with short-chain poly-(R)-3-hydroxybutyrate (cPHB) has been observed, with an average of 21 residues of cPHB per molecule of H-NS. The cPHB, ampiphilic in nature, is thought to play a role in the interaction of H-NS and DNA (Reusch *et al.*, 2002). Figure 1.2 shows a sequence alignment of some selected members of the H-NS family, including paralogous proteins such as Sfh, StpA and Ler.

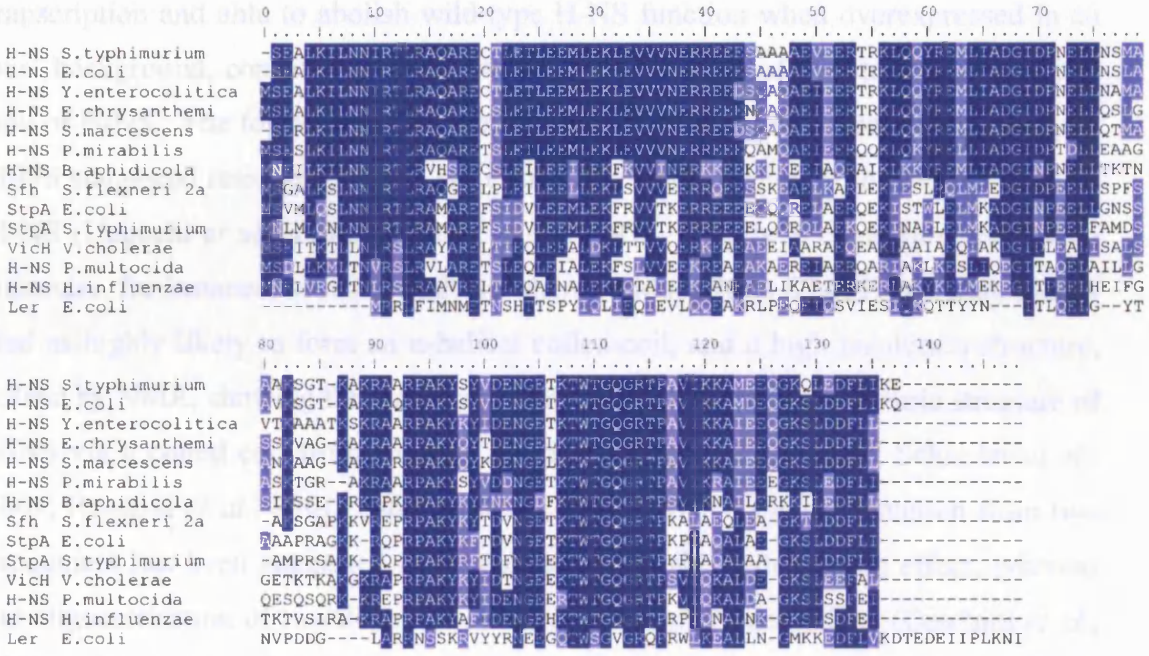


Figure 1.2
Sequence alignment of selected members of the H-NS-like family of proteins. The amino acid residues are numbered with respect to *S. typhimurium* and *E. coli*. Identical residues are coloured dark blue, and similar residues are coloured light blue. Gaps in the alignment are represented with '-'. These proteins were identified by a BLAST 2.0 sequence similarity search on *S. typhimurium* H-NS (Altschul *et al.*, 1997). The proteins were aligned using ClustalW (Thompson *et al.*, 1994). The alignment was graphically enhanced using BioEdit (Hall, 2004).

1.9.1 H-NS consists of two domains, separated by a long linker

H-NS has two structurally independent domains linked by a long, flexible linker. Studies of dominant-negative mutants of H-NS first revealed the two-domain structure of H-NS (Williams *et al.*, 1996). Limited trypsin digestion and phylogenetic analyses supported this domain organisation model of H-NS, and identified an exposed region from residues 77-89 that exhibits the least level of conservation (Cusick and Belfort, 1998). $^{13}C\alpha$ and $^{13}C\beta$ NMR secondary chemical shift data indicate that the residues 60 to 90 do not exhibit any secondary structure either before or after binding DNA (Shindo *et al.*, 1999). In fluorescence energy transfer experiments, a distance of 45 Å between the N-terminal and C-terminal domains of H-NS was measured, suggesting the presence of a long, flexible linker (Schroder *et al.*, 2001).

The use of dominant negative H-NS mutants that are deficient in the ability to repress transcription and able to abolish wild-type H-NS function when overexpressed in an *hns*$^{+}$ background, contributed greatly to the understanding of the structural organisation of H-NS. The formation of heterodimers of N-terminal constructs with wild-type H-NS suggested residues 21 to 63 are critical for the formation of the dimeric core of H-NS (Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997). This is in close agreement with literature, for instance, *in silico* analysis where residues 20 to 60 of H-NS were identified as highly likely to form an α-helical coiled-coil, and a high resolution structure, solved by NMR, showing the stabilisation of an α-helix-rich homodimeric structure of H-NS via a coiled coil from residues 20 to 50 (Smyth *et al.*, 2000; Schroder *et al.*, 2001; Renzoni *et al.*, 2001; Esposito *et al.*, 2002). H-NS dimer formation from two monomers has been suggested to be dominated by the hydrophobic effect, whereas the oligomerisation of two dimers is dominated by polar interactions (Ceschini *et al.*, 2000). The former observation, at least, is in agreement with evidence that H-NS dimerisation is mediated by the formation of a coiled coil (Smyth *et al.*, 2000).

The C-terminal domain of *E. coli* H-NS, corresponding to residues 91 to 137, was initially identified by limited trypsin digestion (Shindo *et al.*, 1995). The domain structure of the C-terminal half of H-NS was later confirmed by analysing NMR $^{13}C\alpha$ and $^{13}C\beta$ chemical shift observations, where no secondary structure was observed between residues 60 to 93 (Shindo *et al.*, 1999).

1.9.2 H-NS C-terminal domain structure

The C-terminal domain of H-NS, corresponding to residues from alanine-91 to glutamine-137, is involved in DNA-binding (Shindo *et al.*, 1995; Ueguchi *et al.*, 1996; Shindo *et al.*, 1999; Badaut *et al.*, 2002). The solution structure of the C-terminal domain of H-NS was solved by NMR, and represents a novel fold, with no homology to any known DNA-binding proteins. The nuclear Overhauser effect data suggested the domain was a relatively flexible and mobile structure. This observation is reflected in the relatively high root-mean-square deviation (RMSD) values between the individual structures and the mean coordinate positions of 72 energy-minimised structures, which were 1.52 ± 0.29 Å and 2.22 ± 0.39 Å for the backbone atoms and the heavy atoms respectively. The averaged structure (in other words the mean coordinate positions of 72 energy-minimised structures) is shown in Figure 1.3. There were four regions of secondary structure identified: from the N-terminus onwards, there is an antiparallel β-sheet (from residues tyrosine-97 to threonine-101 and threonine-106 to tryptophan-109), a α-helix (residues alanine-117 to glutamate-125) and a short 3_{10}-helix (residues leucine-130 to phenylanaline-133) (Shindo *et al.*, 1995).

NMR studies involving changes in chemical shift have identified two regions in the C-terminal domain of H-NS responsible for DNA-binding. Unlike the majority of other DNA-binding proteins, none of the secondary structures were directly involved in binding DNA: instead, two relatively disordered loops (from residues alanine-80 to lysine-96 and from threonine-110 to alanine-117) were involved in interaction with DNA. This lead to the speculation that the flexibility exhibited by the recognition sites may help explain the non-sequence-specific binding of H-NS (Shindo *et al.*, 1999). Analysis of missense mutants of H-NS have suggested two stretches of amino acids in the C-terminal domain, from residues arginine-90 to alanine-95 and threonine-110 to threonine-115 are directly involved in DNA-binding (Ueguchi *et al.*, 1996; Shindo *et al.*, 1999). These observations are in close agreement with fluorescence spectroscopy analysis of the residue tryptophan-108, which has been suggested to be in close proximity to bound DNA (Tippner and Wagner, 1995). Residues 108 to 116 have been suggested to encompass a conserved DNA-binding motif (TWTG-GR-P) (Dorman *et al.*, 1999).

Curiously, N-terminally-truncated H-NS constructs were reported to display oligomerisation: an H-NS mutant displaying residues 60 to 137 and another displaying 70 to 137 were shown to be dimeric and trimeric, respectively (Shindo *et al.*, 1999). Yet research (including the structure of the C-terminal domain of H-NS published by the same group) clearly favours the model whereby the C-terminal domain of H-NS in isolation is monomeric in solution (Shindo *et al.*, 1995; Shindo *et al.*, 1999; Smyth *et al.*, 2000). There is no indication in the literature that the residues 60 to 90 of H-NS could be responsible for oligomerisation in the absence of the N-terminal oligomerisation domain. $^{13}C\alpha$ and $^{13}C\beta$ chemical shift analysis indicate that the residues 60 to 90 do not exhibit any secondary structure either before or after binding DNA (Shindo *et al.*, 1999).

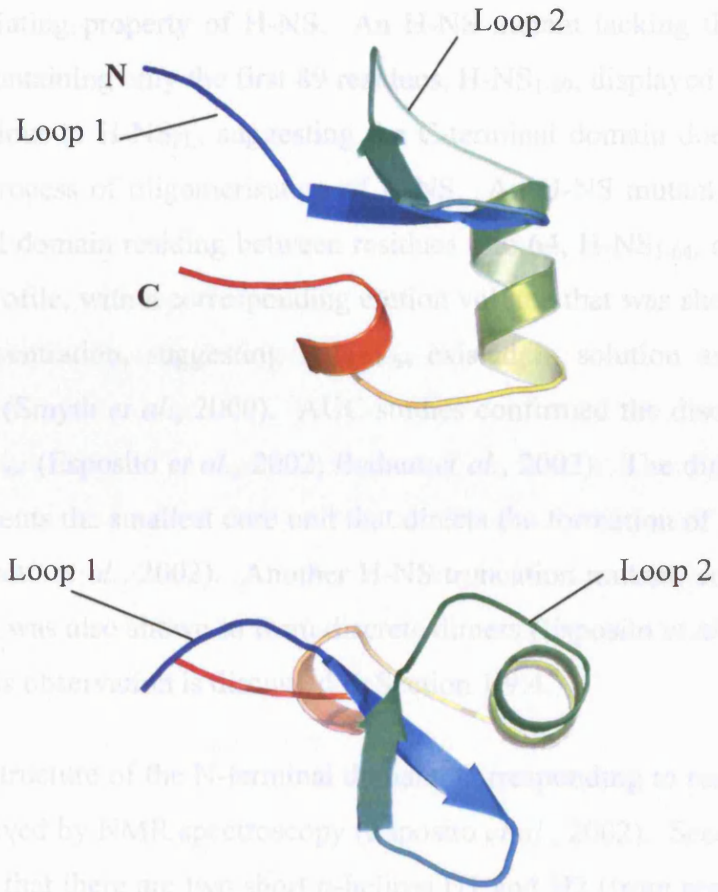


Figure 1.3
Two orthogonal views of the three dimensional structure of the C-terminal domain of *E.coli* H-NS, solved by NMR spectroscopy, corresponding to residues 91 to 137 (Shindo *et al.*, 1995). The structural coordinates file (PDB ID: 1HNR) was obtained from the RCSB PDB. The loop regions suggested to be involved in DNA-binding are labelled 'loop 1' and 'loop 2' (Shindo *et al.*, 1999). These cartoon representations of $H\text{-}NS_{91\text{-}137}$ were created using PyMOL (Delano, 2004).

1.9.3 H-NS N-terminal domain structure

The N-terminal domain of H-NS mediates the higher order oligomerisation, which is crucial for *in vivo* function as a transcriptional repressor (Williams *et al.*, 1996; Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997; Badaut *et al.*, 2002). However, the N-terminal domain of H-NS itself does not interact with DNA directly (Badaut *et al.*, 2002).

Studies of H-NS truncation mutants have aided in the identification of the domain structure of H-NS in the context of the formation of higher order oligomers (Smyth *et al.*, 2000; Esposito *et al.*, 2002). These truncation mutants of H-NS are shown schematically in Figure 1.4. Size exclusion chromatography (SEC) of full-length H-NS resulted in broad, asymmetric peaks with an extended trailing edge, consistent with the self-associating property of H-NS. An H-NS mutant lacking the DNA-binding domain and containing only the first 89 residues, H-$NS_{1\text{-}89}$, displayed similar SEC migration behaviour to H-NS_{FL}, suggesting the C-terminal domain does not contribute towards the process of oligomerisation of H-NS. An H-NS mutant containing only the N-terminal domain residing between residues 1 to 64, H-$NS_{1\text{-}64}$, displayed a symmetric SEC profile, with a corresponding elution volume that was shown to be invariant with concentration, suggesting H-$NS_{1\text{-}64}$ existed in solution as a discrete oligomeric state (Smyth *et al.*, 2000). AUC studies confirmed the discrete dimeric nature of H-$NS_{1\text{-}64}$ (Esposito *et al.*, 2002; Badaut *et al.*, 2002). The dimeric N-terminal domain represents the smallest core unit that directs the formation of higher order oligomers (Esposito *et al.*, 2002). Another H-NS truncation mutant, consisting of residues 12 to 89, was also shown to form discrete dimers (Esposito *et al.*, 2002); the importance of this observation is discussed in Section 1.9.4.

The solution structure of the N-terminal domain, corresponding to residues 1 to 57 of H-NS, was solved by NMR spectroscopy (Esposito *et al.*, 2002). Secondary chemical shifts indicate that there are two short α-helices H1 and H2 (from residues 2 to 7 and residues 10 to 16, respectively) followed by a long α-helix H3 (residues 22 to 49) (Renzoni *et al.*, 2001; Esposito *et al.*, 2002). The H-NS dimer consists of a left-handed parallel coiled-coil (along helix H3), with the two smaller N-terminal α-helices H1 and H2 that fold back onto the helix H3, giving an elongated globular structure. Negatively charged residues dominate the H-NS dimer surface, consistent with the observation that the similarly charged DNA does not interact with the N-

terminal domain. Compared with the C-terminal domain structure of H-NS, the N-terminal domain represents a rigid structure, as shown by the root-mean-square values (RMSD) from an averaged structure of 0.39 ± 0.06 Å and 0.91 ± 0.07 Å for the back-bone and heavy atoms, respectively (Esposito *et al.*, 2002). A representation of the model structure of the N-terminal domain is presented in Figure 1.5.

An alternative model solution structure of the N-terminal domain of H-NS, corresponding to residues 1 to 47, solved by NMR, has been presented (Bloch *et al.*, 2003). The dimeric nature and secondary structure of this model agrees closely with published literature (Renzoni *et al.*, 2001; Esposito *et al.*, 2002). However the two model structures differ significantly in their quaternary structure: the model structure solved by Bloch *et al* has been referred to as a 'handshake', where the two short N-terminal helices of the monomers of H-NS intertwine with each other, and the long helices (corresponding from residues 22 to 46) form an anti-parallel coiled-coil, as shown in Figure 1.6. This model, like the H-NS N-terminal domain structure solved by Esposito *et al*, also represents a relatively rigid structure compared to the C-terminal domain, with root-mean-square deviation (RMSD) values from an averaged structure of 0.51 Å and 1.0 Å for the backbone and heavy atoms, respectively (Bloch *et al.*, 2003). This model structure of H-NS is supported by the recent publication by the same group of the crystal structure of residues 2 to 49 of the paralogue VicH (with 57% sequence identity with H-NS), solved by X-ray crystallography (Cerdan *et al.*, 2003). The N-terminal domain of VicH, as shown in Figure 1.7, shows a 'handshake' fold that is very similar to the H-NS N-terminal structure determined by Bloch *et al.*

The contradiction presented by the two conflicting models of the N-terminal domain of H-NS remains to be resolved. Due to the difficulty in applying NMR techniques to this domain, one of the models may be incorrect due to the following reasons. The backbone amide NH of residues involved in α-helix formation tend to cluster around 8 parts-per-million (ppm) on the ^{1}H chemical shift scale, leading to degeneracy and signal overlap in ^{15}N-edited spectra. Symmetry-related degeneracy leads to ambiguity between inter- and intra-monomer NOE distance restraints. There are a number of clusters of consecutive identical amino acids in the primary sequence of H-NS, leading to potential ambiguity. However, as the model by Esposito *et al* was determined at a temperature of 25 °C, and the model by Bloch *et al* was determined at 35 °C, it is

tempting to speculate that the two structures represent alternative conformations of a temperature-induced rearrangement of the N-terminal domain of H-NS.

A recent study of the involvement of *V. cholerae* H-NS (in other words, VicH) in the transcriptional repression of the ToxR regulon suggested residues 1 to 24 are sufficient for the formation of oligomers in the absence of a coiled-coil domain (Nye and Taylor, 2003). This is in stark contrast to findings based on *E. coli* H-NS, which assert that the formation of the coiled-coil is absolutely essential for the formation of higher order oligomers (Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997), implying there may be some fundamental differences between *E. coli* H-NS and VicH.

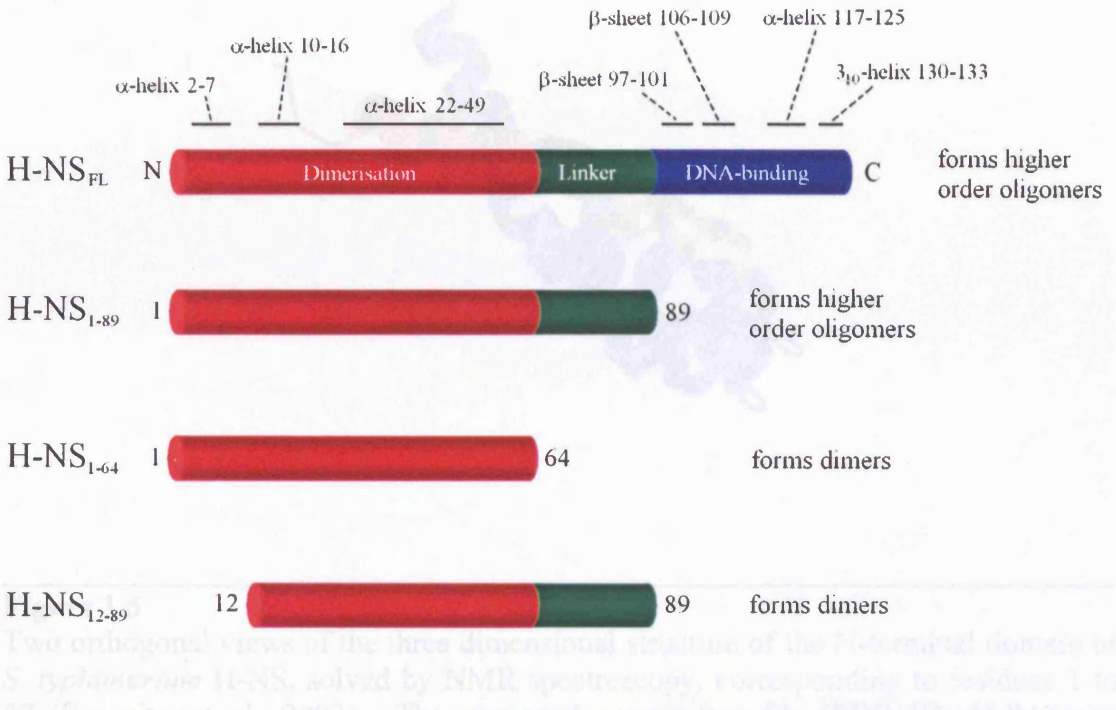


Figure 1.4
Schematic models of the truncation mutants of H-NS displaying differing oligomerisation properties. The N-terminal domain, the linker region and the C-terminal DNA-binding domain are coloured red, green and blue repsectively. The secondary structures are indicated at the top of the figure.

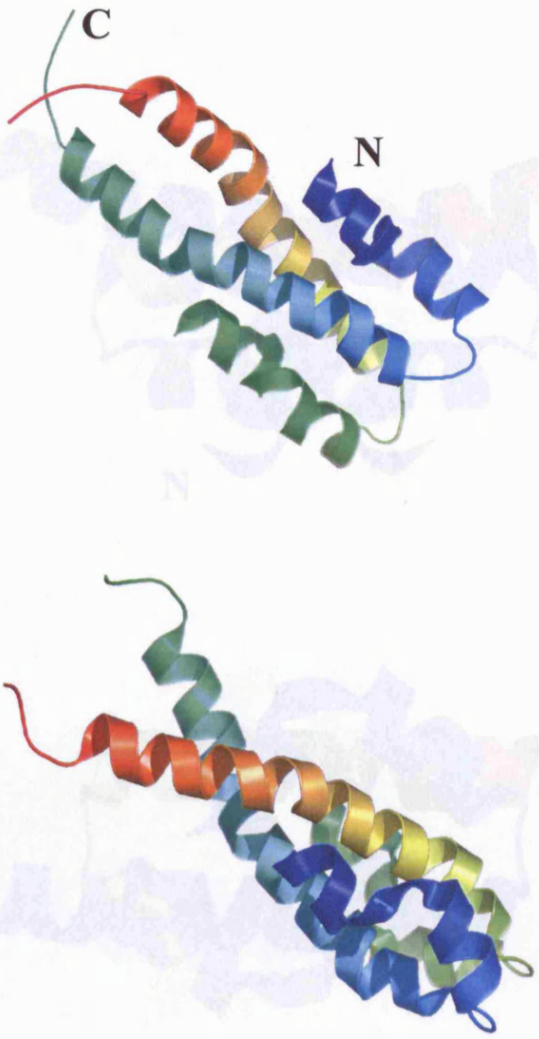


Figure 1.5
Two orthogonal views of the three dimensional structure of the N-terminal domain of *S. typhimurium* H-NS, solved by NMR spectroscopy, corresponding to residues 1 to 57 (Esposito *et al.*, 2002). The structural coordinates file (PDB ID: 1LR1) was obtained from the RCSB PDB. The cartoon representations of the N-terminal domain of H-NS were created using PyMOL (Delano, 2004).

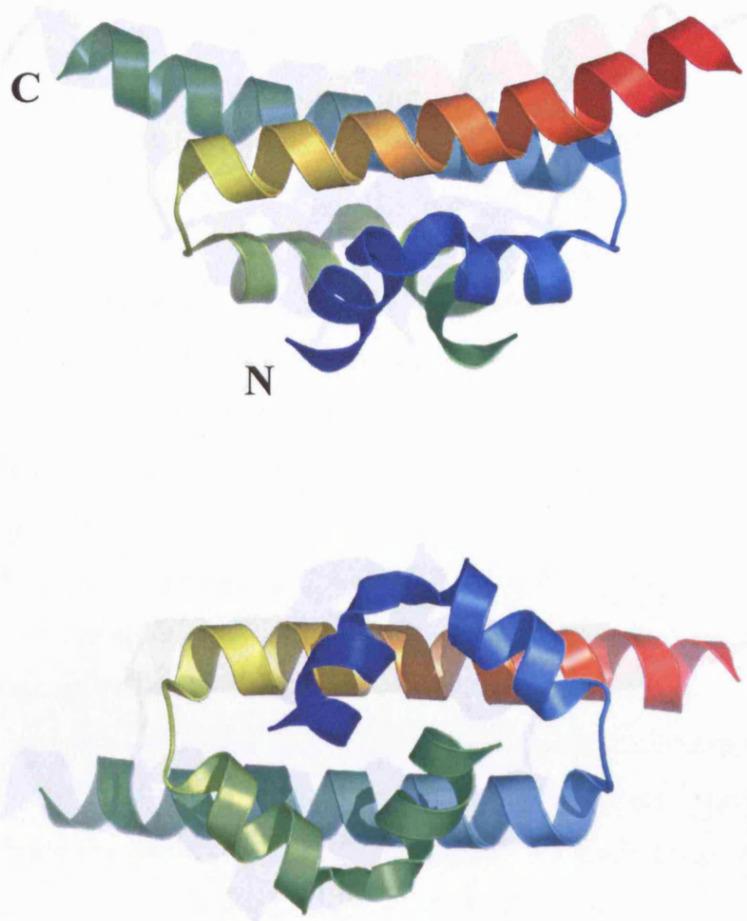


Figure 1.6
Two orthogonal views of the three dimensional structure of the N-terminal domain of *E. coli* H-NS, solved by NMR spectroscopy, corresponding to residues 1 to 46 (Bloch *et al.*, 2003). The structural coordinates file (PDB ID: 1NI8) was obtained from the RCSB PDB. The cartoon representations of the N-terminal domain of H-NS were created using PyMOL (Delano, 2004).

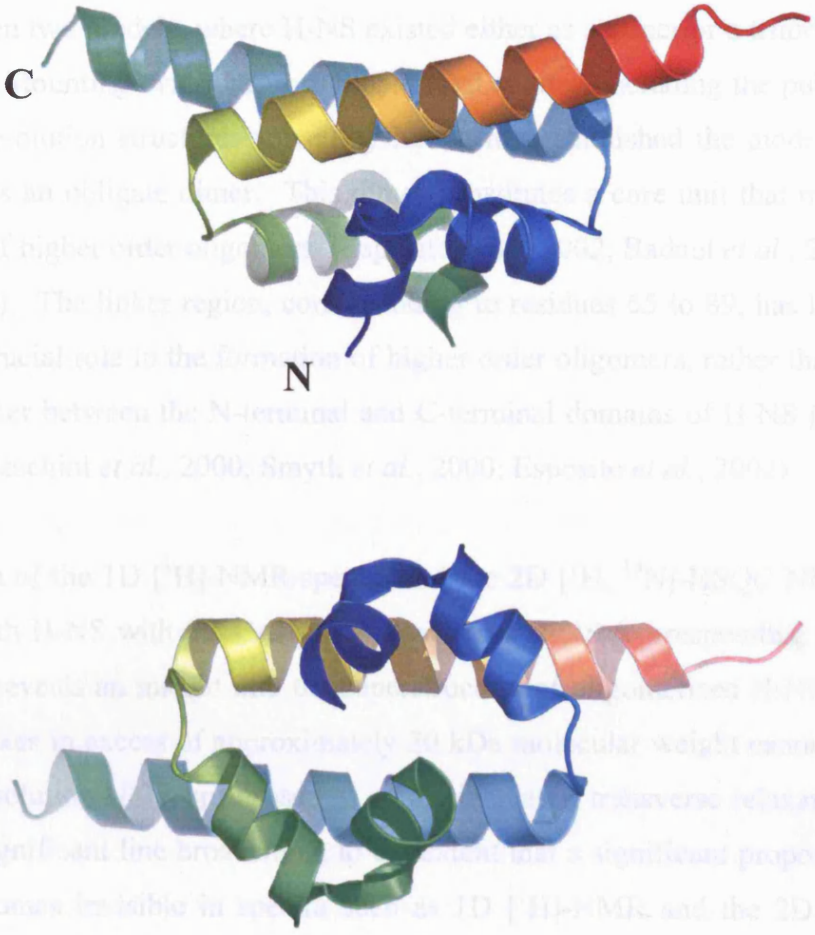


Figure 1.7
Two orthogonal views of the three dimensional structure of the N-terminal domain of *V. cholerae* VicH, solved by X-ray crystallography at 2.5 Å resolution, corresponding to residues 2 to 49 (Bloch *et al.*, 2003). The structural coordinates file (PDB ID: 1OV9) was obtained from the RCSB PDB. The cartoon representations of the N-terminal domain of VicH were created using PyMOL (Delano, 2004).

1.9.4 H-NS forms higher order oligomers – the 'head-to-tail' model

Various studies have attempted to describe the oligomeric form of H-NS. Oligomeric states of H-NS of monomer, dimer, trimer, tetramer and as high as icosamer (20 monomers) have been reported (Ueguchi *et al.*, 1996; Ceschini *et al.*, 2000; Smyth *et al.*, 2000). The nature of the smallest obligate multimer of H-NS was a matter of debate between two models, where H-NS existed either as a dimer or a trimer (Smyth *et al.*, 2000). Mounting evidence from biophysical studies, including the publication of two high-resolution structures solved by NMR, has established the model where H-NS exists as an obligate dimer. This dimer constitutes a core unit that mediates the formation of higher order oligomers (Esposito *et al.*, 2002; Badaut *et al.*, 2002; Bloch *et al.*, 2003). The linker region, corresponding to residues 65 to 89, has been shown to have a crucial role in the formation of higher order oligomers, rather than as a featureless linker between the N-terminal and C-terminal domains of H-NS (Ueguchi *et al.*, 1996; Ceschini *et al.*, 2000; Smyth *et al.*, 2000; Esposito *et al.*, 2002).

Comparison of the 1D [^{1}H]-NMR spectra and the 2D [^{1}H, ^{15}N]-HSQC NMR spectra of full-length H-NS with the C-terminal domain of H-NS (corresponding to residues 90 to 136) reveals an insight into the superstructure of oligomerised H-NS. Proteins and complexes in excess of approximately 30 kDa molecular weight cannot easily be studied by solution NMR spectroscopy, as the increased transverse relaxation rate T_2 results in significant line broadening, to the extent that a significant proportion of the protein becomes invisible in spectra such as 1D [^{1}H]-NMR and the 2D [^{1}H, ^{15}N]-HSQC spectra. The NMR spectra of the full length and the C-terminal domain of H-NS are essentially identical, suggesting that the C-terminal domain attached the full-length H-NS retains similar solution dynamics to the C-terminal domain in isolation, whereas the N-terminal domain exists in a complex with a molecular mass well in excess of 30 kDa. The difference in behaviour of the two domains can be rationalised by their separation by a flexible linker (Smyth *et al.*, 2000).

The most comprehensive model of the mechanism of H-NS oligomerisation is termed the 'head-to-tail' model. The H-NS truncation mutants that support this are shown schematically in Figure 1.4. SEC studies showed that whilst H-NS$_{1\text{-}89}$ was capable of forming higher order oligomers in a fashion similar to wild-type H-NS, H-NS$_{1\text{-}64}$ existed as a discrete dimer. This suggested residues 65 to 89 were crucial for the forma-

tion of higher order oligomers and that the residues 90 to 136 (the DNA-binding domain) do not contribute to higher order oligomerisation of H-NS (Smyth *et al.*, 2000). Subsequent studies involving analysis of the interaction between various truncation mutants of H-NS have identified a second interface on H-NS structure that is crucial for the formation of higher order oligomers (Esposito *et al.*, 2002; Badaut *et al.*, 2002). This second interface is suggested to be located in or around residues 1 to 12 of H-NS. Loss of this second interface near the N-terminus resulted in the loss of the ability to form higher order oligomers, despite the presence of residues 65 to 89, as was observed for the truncation mutant H-NS_{12-89} (Esposito *et al.*, 2002). In the 'head-to-tail' model, higher order oligomers of H-NS are formed by the interaction of the 'head' of one H-NS dimer (found between residues 65 to 89) and the 'tail' of another (found in the N-terminal domain, possibly between residues 1 to 12), in a concentration dependent manner, forming a protein scaffold. The C-terminal domain exhibits higher mobility in solution, with a degree of independence from the scaffold formed by the oligomerised N-terminal domain, constrained only by the long flexible linker. The 'head-to-tail' model is schematically represented in Figure 1.8. An alternative 'head-to-head' model, where the higher order oligomers of H-NS are formed through interaction of the linker region (corresponding to residues 65 to 89), has been ruled out (Esposito *et al.*, 2002). The 'head-to-head' model, where the formation of higher order oligomers is mediated by the residues 65 to 89 of H-NS alone, clearly does not complement the observation that H-NS_{12-89} forms discrete dimers. The 'head-to-tail' model of H-NS oligomerisation is discussed in more detail in Chapter 4. A study of the differential stabilities of mutants of *V. cholerae* H-NS as a result of truncation was suggested to be consistent with the 'head-to-tail' model (Nye and Taylor, 2003).

An alternative model describing H-NS function has been proposed, where both the N- and C-terminal domains of H-NS dimer interact with adjacent regions on a stretch of DNA, resulting in the structuring of the linker region between residues 46-90. This structured linker region can then interact with other similarly structured linker regions of other H-NS dimers bound to DNA, resulting in the bending and condensation of DNA (Bloch *et al.*, 2003).

This model is clearly inconsistent with electrophoretic mobility shift assays that suggest the N-terminal truncation mutant, H-NS_{1-64}, does not bind DNA (Badaut *et al.*, 2002). The model proposed by Bloch *et al.* suggests that the binding of DNA is essential for triggering the structuring of the linker region (46-90), which then allows higher order oligomerisation. However, the formation of higher order oligomers of H-NS (i.e. consisting of multiple units of H-NS dimers) has been observed in the absence of DNA (Ceschini *et al.*, 2000; Smyth *et al.*, 2000; Esposito *et al.*, 2002).

The differential observed DNA-binding characteristics of wild-type H-NS and an R11E/R14A double mutant is crucial to the model proposed by Bloch *et al.*, where the loss of the two basic residues near the N-terminus was suggested to be directly involved in DNA-binding (Bloch *et al.*, 2003). However this observed difference between wild-type H-NS and the R11E/R14A double mutant is consistent with the 'head-to-tail' model of H-NS higher order oligomerisation. The double mutation may result in either the defective formation of the dimeric N-terminal oligomerisation domain, or disrupt local structure in such a way that the 'tail' interface can no longer recognise and bind the 'head' interface of another H-NS dimer. The double mutant can no longer form efficient higher order oligomers, thus is unable to interact with equivalent efficiency to DNA compared to wild-type.

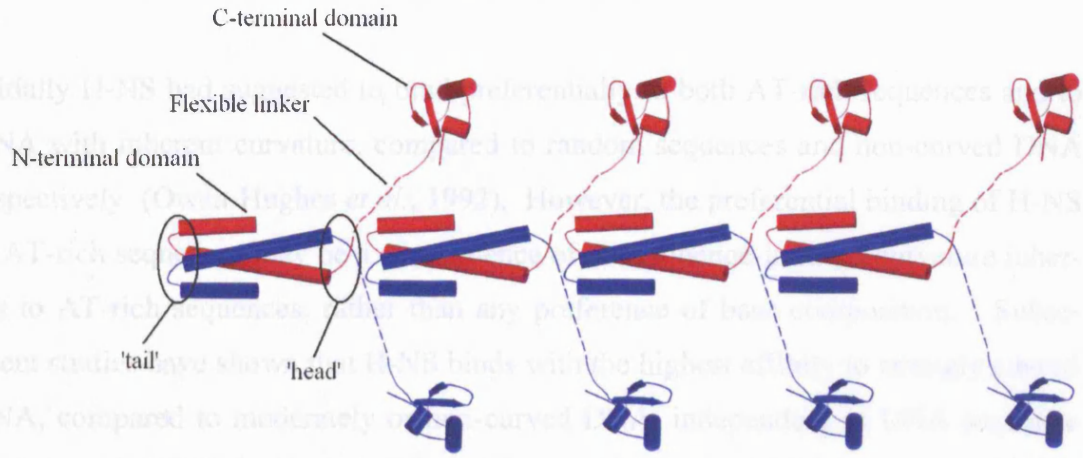


Figure 1.8
A schematic of the 'head-to-tail' model describing the higher order oligomerisation of H-NS. The H-NS dimer is modelled from the NMR structures, with the linker region (residing between residues 65 to 89) represented by a dotted line (Shindo *et al.*, 1995; Esposito *et al.*, 2002). The 'head' region of one H-NS dimer, found between residues 65 to 89, interacts with the 'tail' region of another H-NS dimer, found between residues 1 to 64, forming a higher order oligomer in a concentration-dependent manner. The C-terminal DNA-binding domain is attached to the oligomer by a long flexible linker.

1.9.5 H-NS binds preferentially to curved DNA

H-NS fulfils two seemingly contradictory roles *in vivo*: as a transcriptional repressor at a variety of promoters, and as an architectural protein providing structure to the nucleoid. The former role might require the ability of H-NS to bind specific DNA sequences at or near promoters, whereas the latter role might require H-NS to bind DNA irrespective of sequence. Years of cumulative research has only recently started to allow elucidation of the mechanism of H-NS function; what has become clear, is that H-NS function represents a highly complex phenomenon.

Initially H-NS had suggested to bind preferentially to both AT-rich sequences and to DNA with inherent curvature, compared to random sequences and non-curved DNA respectively (Owen-Hughes *et al.*, 1992). However, the preferential binding of H-NS to AT-rich sequences may be a consequence of the sequence-induced curvature inherent to AT-rich sequences, rather than any preference of base composition. Subsequent studies have shown that H-NS binds with the highest affinity to strongly curved DNA, compared to moderately or non-curved DNA, independent of DNA sequence (Zuber *et al.*, 1994; Jordi *et al.*, 1997; Azam and Ishihama, 1999; Poore and Mobley, 2003). H-NS has been shown to induce further bending of DNA upon binding (Poore and Mobley, 2003). However, the relative affinity of H-NS for curved DNA sequences is less than an order of magnitude greater than to generic, non-curved DNA sequences (Jordi *et al.*, 1997). Furthermore, H-NS has been shown to affect DNA topology *in vivo*, and constrain negative supercoils *in vitro* (Owen-Hughes *et al.*, 1992; Tupper *et al.*, 1994).

The importance of curvature in H-NS/DNA interactions has been highlighted. H-NS mutants with either a proline-115 deletion or proline-115-to-alanine mutation (P115A) displayed a normal affinity to non-curved DNA, but compared to wild type, a reduced capacity to bind curved DNA and ability to actively curve DNA. In contrast to wild-type, the overexpression of these H-NS mutants were shown to not inhibit macromolecular syntheses, non-toxic to cells, and to not cause a drastic compaction of the nucleoid (Spurio *et al.*, 1997). Fluorescence spectroscopy studies have shown two distinguishable H-NS/DNA complexes, depending on the curvature exhibited by the DNA sequences. Interaction of H-NS with curved DNA sequences is

characterised by a hydrophobic interaction, whereas with non-curved DNA, the interaction is exclusively electrostatic (Tippner and Wagner, 1995).

It is worth noting that there are curved DNA sequences near most promoters, including promoters that are not under the control of H-NS (Ohyama, 2001). H-NS specificity for target promoters cannot be explained by DNA curvature alone (Jordi *et al.*, 1997). This implies that, in certain operons, other unidentified auxiliary proteins may be required for full H-NS repression.

Force-extension measurement experiments involving λ-phage DNA at 25°C have shown H-NS to densely decorate DNA molecules at a ratio of 15 to 20 base pairs of DNA to one dimer of H-NS, to the extent that the H-NS/DNA complex displays a higher degree of bending rigidity compared to naked DNA. The suggestion that H-NS oligomerises along the DNA molecule conforms to the observation that the DNA is protected from the restriction nuclease *Sau*3A. At 37°C or higher osmolarity however, H-NS appears neither to exert any effect on the bending rigidity of DNA nor to protect the DNA from *Sau*3A restriction nuclease, implying H-NS is no longer bound to the DNA (Amit *et al.*, 2003). This study opposes the current opinion that H-NS induces DNA compaction: rather, interaction with H-NS forces the DNA into an extended, rigid structure. However, this study should be interpreted with caution: force extension measurements were made with individual λ-phage DNA molecules in isolation, as opposed bulk measurements. Consequently the effective concentration of H-NS may be much higher than those cited in other studies, and the λ-phage DNA may be saturated with H-NS, leading to the observed effects on persistence length and rigidity of λ-phage DNA (Dame and Wuite, 2003). Indeed, other studies have observed differences between specific and non-specific binding of H-NS to DNA (also see Section 1.9.7) (Tippner and Wagner, 1995; Spurio *et al.*, 1997).

Analysing the interaction between H-NS and DNA has been made difficult by factors such as the lack of a defined DNA sequence to which H-NS binds, and the lack of a clear mechanism of DNA-binding by H-NS. Nonetheless, gel retardation studies have provided clues: an apparent K_d of 250 nM and 1 mM were estimated for H-NS and H-$NS_{90\text{-}136}$ respectively, for an oligonucleotide displaying a repeat sequence $d(GGCAAAAAC)_{12}$ (Shindo *et al.*, 1995). Studies involving protection of a synthetic

5A6A *gal* promoter by H-NS led to an estimate of the K_d of ~5 nM (Rimsky *et al.*, 2001). These studies suggest high H-NS affinities for some specific DNA sequences.

1.9.6 Mechanism of function

The ability of H-NS to bind DNA is insufficient to explain H-NS function: the formation of higher order oligomers of H-NS is crucial to its *in vivo* function as a transcriptional repressor (Williams *et al.*, 1996; Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997; Badaut *et al.*, 2002). This is consistent with the observation that the C-terminal domain of H-NS in isolation binds curved-DNA at least three orders of magnitude weaker than wild-type H-NS (competent at forming higher order oligomers) (Shindo *et al.*, 1995).

H-NS can bind and repress transcription from specific promoters, whilst at the same time exhibiting only weak specificity for DNA sequences with inherent curvature. Clearly the mechanism of H-NS function differs from 'classical' DNA binding proteins (Ueguchi and Mizuno, 1993; Ussery *et al.*, 1994). Initially, it was suggested that H-NS affects transcription by affecting the degree of DNA supercoiling, where the change in DNA topology interferes with the ability of transcription factors to bind DNA (Tupper *et al.*, 1994; Jordi *et al.*, 1995). In another model, termed 'transcriptional silencing', H-NS binds and polymerises along a stretch of DNA, thus sterically hindering RNA polymerases from binding relevant gene promoters (Goransson *et al.*, 1990; Williams and Rimsky, 1997). The importance of other factors in H-NS function, such as temperature, osmolarity, growth phase and pH, has been highlighted: for example, H-NS was shown to repress the transcription of *virF* by binding cooperatively two distally separated sites upstream of the promoter (at positions -1 and -250) below 32°C, but not at 37°C (Falconi *et al.*, 1998a). Although a few genes found to be under positive regulation by H-NS have been reported (Johansson *et al.*, 1998; Starcic-Erjavec *et al.*, 2003), there is not yet any evidence for H-NS acting directly as a transcriptional activator (Dorman, 2004), suggesting the repressive effect of H-NS on transcription is universal.

The *E. coli proU* operon encodes a transport system for the osmoprotectant glycine betaine, and its expression is induced in response to osmotic stress. The *proU* operon has been studied extensively as a model H-NS dependent promoter system. No se-

quence-specific regulatory protein has been implicated in *proU* regulation: only the *trans*-acting repressor H-NS, the *cis*-acting downstream regulatory element (DRE) and a region of curved DNA (to which H-NS has exhibited preferential binding) approximately 200 bp upstream of the promoter are involved (Jordi *et al.*, 1995; Jordi and Higgins, 2000; Rajkumari and Gowrishankar, 2001), although it has been suggested that StpA may play a role at the *proU* promoter (Free *et al.*, 2001). The differential effects of C-terminal and N-terminal dominant-negative mutants of H-NS at two model promoters of *proU* and *5A6AgalP1* suggest that there are at least two different mechanisms of promoter repression by H-NS, where *proU* promoter repression requires the formation of higher order oligomers by H-NS, and *5A6AgalP1* promoter repression requires recognition of curved DNA sequences by H-NS (Williams *et al.*, 1996).

In a model termed 'nucleation', consistent with studies involving DNase I footprinting and transcription modulation by H-NS on semi-synthetic *gal*P1 promoters, inhibition of transcription has been suggested to occur by two distinct mechanisms. In this model, H-NS 'nucleates' at specific sites in the DNA at low concentrations, allowing an increase in H-NS occupancy at lower affinity sites, in turn leading to the formation of a complex that prevents RNA polymerase binding at the promoter. At higher H-NS concentrations, H-NS binds and polymerises along the DNA in a non-specific manner, effecting repression (Rimsky *et al.*, 2001; Badaut *et al.*, 2002). Indeed, there is speculation that these two different modes of DNA-binding reflect the multiple functions of H-NS *in vivo*: the 'nucleation' model, involving interaction with specific DNA, may mediate the regulatory functions of H-NS, whereas the DNA-binding to non-specific sequences may play a role in the general structuring of the nucleoid (Badaut *et al.*, 2002).

Clearly, another important consideration of H-NS mechanism *in vivo* is the involvement of other proteins, leading either to the formation of specific nucleoprotein complexes fully competent at repression, or antagonism of the repressive effect of H-NS via a transcriptional activator, that allow specific modulation of gene expression (also see Section 1.8). For instance, the antagonistic effects of H-NS and FIS mediate the thermoregulation of the expression of VirF, a transcriptional activator of the invasion

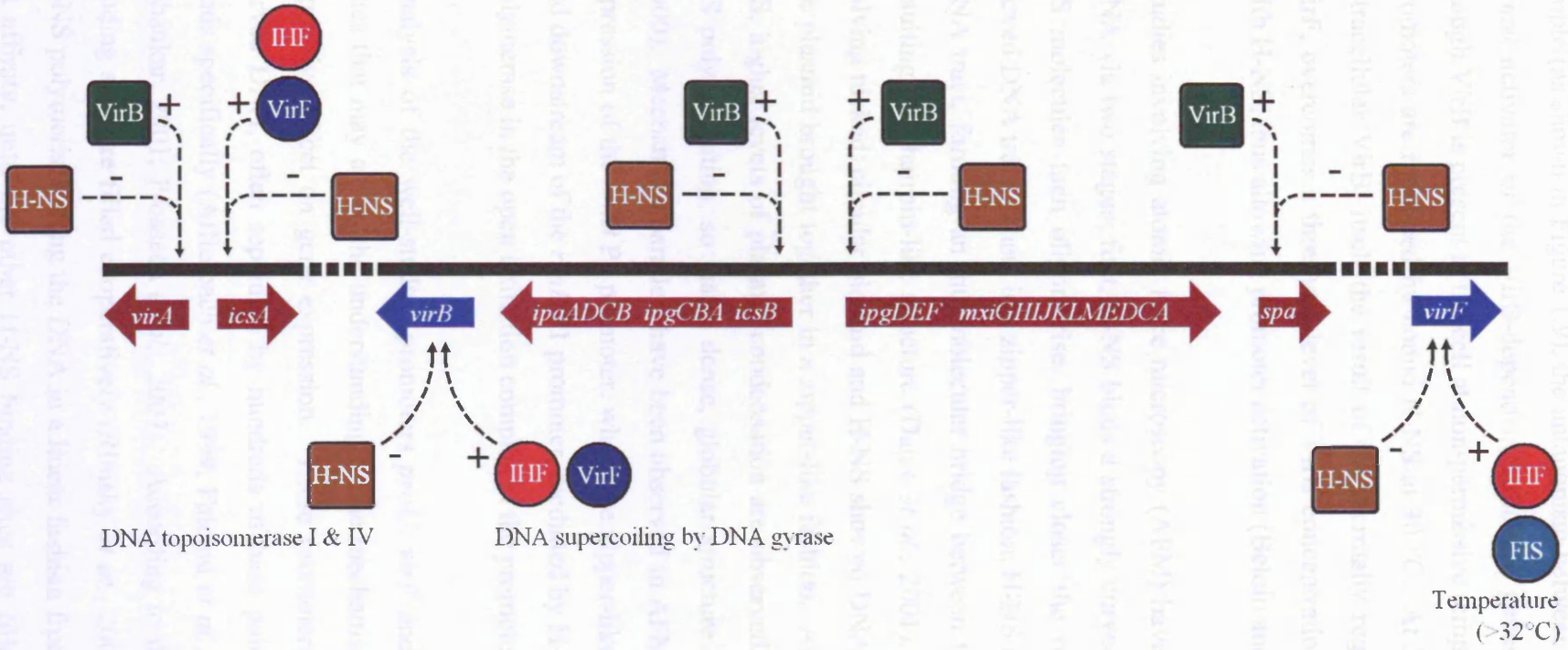


Figure 1.9
A schematic showing some of the key features of the transcriptional regulation of the virulence genes of *S. flexneri*. The regulatory genes *virF* and *virB* are shown in blue, and the structural virulence genes are shown in dark red. H-NS and VirB are represented as brown and green squares respectively. VirF, IHF and FIS are represented as dark blue, red and light blue circles respectively. The positive (+) and negative (-) regulatory effects of these proteins on transcription are shown schematically.

Thermoregulation of *virF* is mediated by H-NS, IHF and FIS, resulting in increased levels of the gene product VirF above 32°C (Prosseda *et al.*, 2004). The VirF is then activates transcription of genes such as *virB* and *icsA*, overcoming the repressive effect of H-NS. VirB then activates the expression of the VirB-dependent structural virulence genes, once the intracellular amount of VirB exceeds a threshold level (Dorman and Porter, 1998). Every step of the regulatory cascade of these virulence genes are repressed by H-NS (Beloin and Dorman, 2003). This figure was adapted from Dorman and Porter, 2003.

functions of *S. flexneri* as shown in Figure 1.9 (Prosseda *et al.*, 2004). In another example (as shown in Figure 1.9), the antagonism between H-NS and VirB, a transcriptional activator of the VirB-dependent structural genes, has been reported. Even though VirB is present in the cell at non-permissive temperatures, the VirB-dependent promoters are repressed by bound H-NS at 30 °C. At 37 °C, the increased level of intracellular VirB, itself the result of the thermally regulated activation of *virB* by VirF, overcomes a threshold level of VirB concentration that may directly compete with H-NS, thus allowing promoter activation (Beloin and Dorman, 2003).

Studies involving atomic force microscopy (AFM) have suggested H-NS condenses DNA via two stages: first, H-NS binds a strongly curved DNA tract. The bound H-NS molecules then oligomerise, bringing closer the two DNA arms flanking the curved DNA tract. Then, in a zipper-like fashion, H-NS oligomerises from the curved DNA tract, forming an intramolecular bridge between the two flanking DNA arms, resulting in a hairpin-like structure (Dame *et al.*, 2001). A previous AFM study involving nicked, circular plasmid and H-NS showed DNA strands on opposite sides of the plasmid brought together in a zipper-like fashion. At higher concentrations of H-NS, higher levels of plasmid condensation are observed, perhaps through further H-NS polymerisation, so that a dense, globular structure was observed (Dame *et al.*, 2000). Mechanistic parallels have been observed in AFM studies of H-NS-dependent repression of the *rrnB* P1 promoter, where the zipper-like structure of DNA upstream and downstream of the *rrnB* P1 promoter, mediated by H-NS, was shown to trap RNA polymerase in the open intitiation complex at the promoter (Dame *et al.*, 2002).

Analysis of the well-studied promoters *proU*, *virF* and *rrnB* P1 reveal interesting clues that may aid in the understanding of the mechanism by which H-NS exerts its repressive effect on gene expression. These promoters display several regions of curved DNA, often separated by hundreds of base pairs of DNA, to which H-NS binds specifically (Afflerbach *et al.*, 1999; Falconi *et al.*, 2001; Rajkumari and Gowrishankar, 2001; Prosseda *et al.*, 2004). According to the 'nucleation' model, these binding sites are filled cooperatively (Rimsky *et al.*, 2001). This implies that either H-NS polymerises along the DNA in a linear fashion from the binding site with highest affinity, until the other H-NS binding sites are filled, or that H-NS molecules bound in the high affinity sites somehow stabilise H-NS binding at the other affinity

sites through space. The latter proposal is possible if the DNA molecule bends in a way so that the two H-NS-binding sites are brought together in close proximity, allowing H-NS molecules to bind both sites cooperatively and oligomerise. This is consistent with the observation that the *virF* promoter has a strong bend between the H-NS binding sites that brings them together through space (Prosseda *et al.*, 2004), and with an AFM study of the *rrnB* P1 promoter, which showed the DNA regions flanking the *rrnB* P1 promoter bridged by H-NS, resulting in a repressive hairpin-like structure (Dame *et al.*, 2002). Indeed, H-NS has been shown to form hairpin-like structures by lateral condensation of generic DNA (Dame *et al.*, 2000; Dame *et al.*, 2001). The formation of this hairpin-like structure mediated by the 'bridging' of lateral stretches of DNA may form the basis of a general mechanism of repression of gene expression by H-NS (Dorman, 2004). A schematic model describing the hairpin-like complex through which H-NS exerts its repressive effects on the *virF* promoter is shown in Figure 1.10. However, promoters may differ in other details, for instance, the requirement for different activator proteins. Transcription from *hns* and *virF* requires FIS (Falconi *et al.*, 1996; Falconi *et al.*, 2001), whereas transcription of *virB* requires VirF (Beloin and Dorman, 2003). Another example of a promoter that fits the model of repression by the formation of an H-NS/DNA hairpin-like complex is the type III secretion genes of enteropathogenic *E. coli* (EPEC). The *LEE2* and

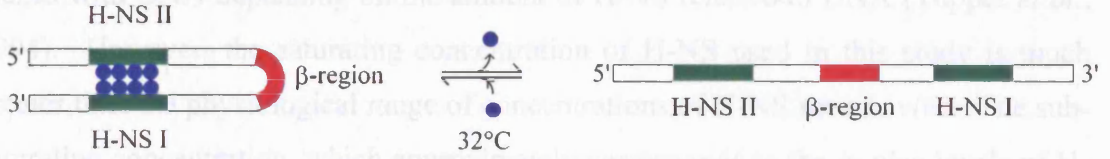


Figure 1.10
H-NS is represented by blue spheres, the H-NS binding sites are coloured green and the β-region of DNA whose conformation is sensitive to temperature is coloured red. At non-permissive temperatures, the β-region of DNA, also referred to as a 'temperature-sensitive hinge', is bent, thus allowing the two H-NS binding sites to be bridged together by H-NS. At temperatures above 32°C, the β-region undergoes a conformational change and straightens out, resulting in the release of H-NS from the repressive complex (Falconi *et al.*, 1998a; Prosseda *et al.*, 2004).

LEE3 operons are expressed from two divergent overlapping promoters, which require H-NS for repression and Ler for activation. Two AT-rich H-NS binding sites were identified upstream and downstream of the promoter sequences, termed silencer regulatory sequences (SRS), both of which were shown to be crucial for repression under non-permissive conditions (Bustamante *et al.*, 2001).

A recent study of the cryptic *bgl* operon revealed another possible mechanism for H-NS-dependent repression of gene expression. In addition to the binding of H-NS to a region of AT-rich DNA upstream of the promoter to silence *bgl* by occlusion of RNA polymerase, another region of curved DNA downstream of the promoter was found to contribute towards repression of *bgl*. It was suggested that H-NS acts as a 'road-block' to RNA polymerase, thereby preventing elongation during transcription (Dole *et al.*, 2004).

1.9.7 The amount of HNS relative to DNA is important

Many studies have suggested that there are two different modes of H-NS binding to DNA, resulting in specific and non-specific interactions; these interactions appear to be determined by the curvature displayed by the target DNA sequence and the relative amounts of H-NS and DNA. For example, whilst H-NS efficiently constrains negative supercoils at sub-saturating concentrations, at saturating concentrations, DNA supercoils are not constrained by H-NS, suggesting H-NS forms two distinct complexes with DNA depending on the amount of H-NS relative to DNA (Tupper *et al.*, 1994). However, the saturating concentration of H-NS used in this study is much greater than the physiological range of concentrations of H-NS seen *in vivo*. The sub-saturating concentration, which approximately corresponds to the *in vivo* levels pf H-NS, is likely to be more significant physiologically (Tupper *et al.*, 1994). As the abundance of H-NS has been estimated at one molecule of H-NS per 400 bp of chromosomal DNA, saturating levels of H-NS are unlikely to reached *in vivo* (Ussery *et al.*, 1994). Fluorescence spectroscopy studies have shown differences in H-NS-binding with specific and non-specific DNA, i.e. curved and non-curved DNA sequences respectively. The former is characterised by a hydrophobic interaction, whereas the latter exhibits electrostatic properties (Tippner and Wagner, 1995). Results from a study involving force extension measurements of H-NS/λDNA complexes contradict current opinions on H-NS function. The inconsistency may in part

be explicable by differential modes of H-NS binding to DNA, itself a consequence of differences in saturation of DNA by H-NS (Amit *et al.*, 2003; Dame and Wuite, 2003). An *in silico* analysis of the *E. coli* genome suggested that there is only sufficient H-NS in the cell to occupy 50% of predicted bends, or high-affinity H-NS binding sites, in the whole genome. As the amount of H-NS roughly matches the number of specific H-NS binding sites, non-specific binding was suggested to play only a minor role *in vivo* (Atlung and Ingmer, 1997).

1.9.8 DNA plays a crucial role in mediating cellular responses to changes in environment

The molecular mechanism of modulation of transcription of an H-NS-dependent gene cannot be explained by H-NS function alone: DNA plays a crucial role in mechanisms responsive to the environment. It has long been established that DNA structure is a highly dynamic entity: environmental factors such as salt shock, high osmolarity, anaerobic shock, anaerobiosis, nutrient shifts, growth phase and changes in temperature have been shown to influence DNA topology (Drlica, 1992; Karem and Foster, 1993). In a study of the modulation of *aniG* expression by a repressor EarA, reminiscent of promoters repressed by H-NS, a change in expression of *aniG* was observed in response to aerobiosis, low osmolarity, low pH, a mutation in *topA* and the presence of novobiocin. It was suggested that these diverse factors were unlikely to all affect expression by modifying the repressor EarA at the protein level. Instead, a model was proposed where modulation of *aniG* expression depended upon a 'dynamic interplay' between repressor affinity for the promoter, which depended upon DNA topology, and promoter affinity for the repressor, which depended on factors such as pH and osmolarity that affect DNA topology (Karem and Foster, 1993).

Parallels may be drawn between the modulation of *aniG* expression and repression of H-NS-dependent genes: the majority of H-NS-dependent genes are responsive to environmental cues such as changes in pH, temperature, osmolarity and growth phase that affect DNA topology. The downstream regulatory element (DRE) of the *proU* promoter has been suggested to act as an 'environmental sensor', where changes in structure induced by changes in the environment influence H-NS-dependent repression (Ueguchi and Mizuno, 1993; Jordi *et al.*, 1997; Jordi and Higgins, 2000). The regulation of *proU* by H-NS has been referred to as 'conditional', where the repres-

sive effect of H-NS was shown to depend on the topology of the target DNA (Ueguchi and Mizuno, 1993). The expression of the H-NS-dependent *flhDC* master operon could be modulated by the alteration of DNA topology. For example, *flhDC* expression could be affected by many factors such as the presence of novobiocin (a DNA gyrase inhibitor), the overexpression/deletion of DNA gyrase and the use of linearised DNA instead of supercoiled DNA. Also, overexpression of a subunit of DNA gyrase (*gyrB*) allowed partial restoration of *flhDC* modulation in a *hns* mutant (Soutourina *et al.*, 2002). The thermoregulation of VirF, a transcriptional activator of the invasion functions of *S. flexneri*, is resolved by the antagonism of H-NS and FIS at the *virF* promoter, which contains a curved segment of DNA, with the apex of which is located halfway between two H-NS binding sites (Falconi *et al.*, 2001; Prosseda *et al.*, 2004). Two FIS binding sites are located near and on the DNA curve apex, sandwiched between the H-NS binding sites (Prosseda *et al.*, 2004). Changes in electrophoretic mobilities of *virF* promoter fragments suggested that this DNA sequence displays a specific temperature-dependent structural transition at around 32°C, in a fashion independent of H-NS, and has been referred to as a 'temperature-sensitive hinge' (Falconi *et al.*, 1998b). The temperature increases were also shown to shift the location of the curve apex away from the FIS binding sites (Prosseda *et al.*, 2004). There was no H-NS-dependent change in DNA topology in the *virF*-carrying plasmid, and the aggregation state of H-NS was not influenced by temperature in the critical range of *virF* regulation (30°C~34°C), suggesting a more passive role for H-NS in the repression of *virF* (Falconi *et al.*, 1998b). A model was proposed whereby at non-permissive temperatures (less than 32°C), the curvature of the DNA bring the two H-NS binding sites close together, as shown in Figure 1.10. H-NS then binds the sites and oligomerise, thus creating intramolecular bridge, resulting in *virF* repression. Temperature-induced changes in DNA conformation at permissive temperatures result in relaxation of the repressive H-NS/DNA complex, whilst allowing FIS to antagonise the repressive effect of H-NS (Prosseda *et al.*, 2004). The thermoregulation of the VirB-dependent structural genes of *S. flexneri* involves the antagonism between the repressive effect of H-NS and an activator protein, VirB. Interestingly, thermoregulation of the VirB-dependent genes is still observed in a *virB hns* double mutant (albeit reversed in response compared to wild-type), suggesting a more fundamental temperature-response system, possibly at the level of temperature-dependent DNA topology (Beloin and Dorman, 2003). Thus changes in environment may cause alterations

in DNA conformation that could be the basis of differential H-NS-protein interactions, which lead to regulation that is sensitive to different conditions.

1.10 Overview of StpA

StpA (suppression of the *td*⁻ phenotype) from *E. coli* was first identified by its ability to suppress the mutant phenotype of the *td*⁻ gene from bacteriophage T4. StpA has 134 amino acids and shares 58% sequence identity (or 67% sequence similarity) with the paralogous gene H-NS (Zhang and Belfort, 1992; Zhang *et al.*, 1996). Studies of *stpA* mutants have yet to identify a phenotype (Sonden and Uhlin, 1996); hence a clear role for StpA has been difficult to define. However, research suggests that StpA function is intimately linked with H-NS. StpA and H-NS share many common properties. For example, both proteins have been shown to have similar structures and domain organisation (Williams *et al.*, 1996).

1.11 A definite role for StpA has yet to be defined

The upregulation of StpA in an *hns* mutant has been shown to substitute for H-NS and control a subset of genes normally regulated by H-NS, leading to the suggestion that StpA functions as a molecular back-up of H-NS (Sonden and Uhlin, 1996). It should be noted that some genes were shown to be under the control of H-NS alone, whereas none were shown to be under the control of StpA alone (Zhang *et al.*, 1996). The derepression of the *adi* gene encoding arginine decarboxylase as a result of an *hns* mutation can be complemented by plasmids expressing StpA (Shi and Bennett, 1994). However, the speculation that StpA is a molecular back-up for H-NS has been largely dismissed, one reason being that the loss of control of a large number of H-NS-dependent genes resulting from an *hns* mutation is not recovered by increased levels of StpA: a study conducted using DNA arrays, tracking the level of several gene transcripts involved in resistance to low pH and motility in *hns* or *stpA* mutants or an *hns stpA* double mutant, found that StpA was not involved in either of these two H-NS related phenotypes (Bertin *et al.*, 2001). On the other hand, the levels of StpA, upregulated as a result of an *hns* mutation, only reach about 10% of the amount of H-NS observed in a wild-type strain. It has been suggested that StpA does not compensate for an *hns* mutation due to the lower levels of expression of StpA, rather than a specific difference in activity at promoters (Sonnenfield *et al.*, 2001).

A possible *in vivo* role for StpA, distinct from H-NS, has been identified in *ompF* porin expression, by modulation of *micF* mRNA stability. The *micF* gene encodes a small, untranslated RNA molecule of 93 nucleotides that post-transcriptionally regulates *ompF* expression by binding *ompF* mRNA and inhibiting transcription and inducing mRNA degradation. H-NS and StpA both indirectly regulate *ompF* expression: H-NS represses transcription of *micF*, whilst StpA enhances degradation of the *micF* mRNA (Deighan *et al.*, 2000; Delihas and Forst, 2001).

The *E. coli* cryptic *bgl* operon is involved in the uptake and utilisation of β-glucoside sugars. H-NS is required, but not sufficient, for *in vitro bgl* repression. Studies have shown that cells harbouring mutant H-NS fragments containing only the N-terminal oligomerisation domain retain a degree of repression due to the presence of StpA, which is thought to function as a molecular adapter. Similar effects have been observed at other promoters such as *stpA* and *proU*, suggesting a wider role for StpA at H-NS-dependent promoters (Free *et al.*, 1998; Free *et al.*, 2001). Another study has shown the sigma factor RpoS to be crucial for *bgl* repression (Ohta *et al.*, 1999).

1.12 Regulation of StpA

The regulation of expression of StpA can vary in response to a variety of conditions, unlike H-NS, whose levels remain relatively constant in response to most environmental stimuli. The genes *hns* and *stpA* exert parallel autogenous control and cross-regulation, with H-NS assuming the dominant cellular role (Zhang *et al.*, 1996; Sonden and Uhlin, 1996). The expression of *stpA* was shown to be *trans*-activated in minimal medium in a leucine-responsive regulatory protein-dependent (Lrp) manner (Sonden and Uhlin, 1996; Free and Dorman, 1997). Approximately twice the amount of *stpA* transcript was observed at 37°C compared to 26°C (Sonden and Uhlin, 1996). The *stpA* gene was found to be highly transiently induced in the mid-exponential phase during growth in rich medium, induced by stresses such as osmotic shock and changes in temperature whilst in rich medium, and repressed during carbon-starvation in minimal medium. Transcription from the *stpA* promoter was shown to be severely repressed by H-NS, leading to little StpA in the cell (Zhang *et al.*, 1996; Free and Dorman, 1997). However this is in disagreement with results from a quantitative Western blotting experiment in which the stationary-phase levels of StpA were sug-

gested to be around 10,000 molecules per cell, almost twice the amount of H-NS in the cell (Ali Azam *et al.*, 1999).

1.13 Structural aspects of StpA

The domain organisation of StpA has been shown to have a high degree of similarity with that of H-NS; limited trypsin digestion profiles of StpA closely match that of H-NS, and phylogenetic analysis of the amino acid sequence reveal the two-domain structure of StpA, with an unconserved solvent-exposed linker region between residues 77 to 90 (Cusick and Belfort, 1998). Whilst StpA and the C-terminal domain of StpA (corresponding to residues 81 to 134) exhibited *trans*-splicing of a mutant *td* intron, the N-terminal domain (corresponding to residues 1 to 80) exhibited no such activity (Cusick and Belfort, 1998). However a high-resolution three-dimensional structure of StpA has yet to be published.

StpA has been shown to form higher-order oligomers *in vitro*, including tetramers (Johansson *et al.*, 2001). StpA was found to have a rapid turnover in an *hns*$^-$ background due to proteolysis by Lon protease, whereas the reverse situation did not occur. A single amino acid change resulted in a Lon protease-resistant mutant of StpA, suggesting StpA is present in *E. coli* in a heteromeric form with H-NS (Johansson and Uhlin, 1999). The formation of heterooligomers of StpA with a variety of H-NS mutants has been shown to confer resistance to proteolysis of StpA by Lon, suggesting H-NS interacts with StpA at two distinct locations on H-NS (Johansson *et al.*, 2001).

1.14 DNA-binding properties of StpA

StpA does not show any sequence specificity in DNA-binding (Azam and Ishihama, 1999; Sonnenfield *et al.*, 2001), although StpA has been shown to exhibit a preference for curved DNA sequences (Sonnenfield *et al.*, 2001). StpA has a four- to six-fold higher affinity for DNA compared to H-NS (Sonnenfield *et al.*, 2001). StpA has been shown to exhibit parallels with H-NS in its DNA-binding properties. Overexpression of both H-NS and StpA inhibited transcription from synthetic *gal* promoter with a curved sequence upstream. The products of DNAse I treatment of DNA with either bound H-NS or StpA are consistent with the proposition that both proteins preferentially bind curved DNA sequences. Both proteins have been shown to constrain supercoils in *vitro* (Zhang *et al.*, 1996).

1.15 RNA chaperone activity of StpA: is it relevant?

StpA has been shown to display superior RNA chaperone activity, both *in vivo* and *in vitro*, compared to H-NS, apparently by the ability of StpA to promote association and dissociation of RNA (Zhang *et al.*, 1996). StpA bound to RNA non-specifically, and was shown not to be required for the catalytic step itself (Zhang *et al.*, 1995). StpA was shown to promote the efficient splicing of the *td* group I intron pre-mRNA, by loosening tertiary structure of misfolded RNA and allowing the RNA to refold into native conformations (Waldsich *et al.*, 2002; Mayer *et al.*, 2002). However the *in vivo* relevance for the RNA chaperone activity has yet to be established.

1.16 Concluding remarks

The members of the nucleoid-associated H-NS family of proteins provide insight into the cellular machinery of scientifically and medically important Proteobacteria. H-NS has been shown to affect the bacterial cell at many different levels, from the organelle structure of the nucleoid to global gene expression and the ability to adapt to a wide range of environments. H-NS is intimately linked with pathogenicity exhibited by many bacteria, and further research may have implications in the prevention and treatment of diseases caused by these facultative organisms.

Comparison of H-NS and StpA presents a challenge in the understanding of the functional mechanisms of H-NS-like proteins. Whilst the current opinion is not clear about the relative significance of the two proteins, they are biologically relevant. Research described in this thesis is biophysical in nature, concentrating on techniques such as NMR, analytical gel filtration chromatography, isothermal titration calorimetry (ITC) and circular dichroism spectroscopy (CD), to address the issues concerning the function of H-NS and its paralogous protein StpA. In Chapter 3 the determination by NMR of the solution structure of the C-terminal domain of StpA is presented, and the resultant structure compared with the structure of the C-terminal domain of H-NS. The application of a series of experiments to H-NS and StpA is described in Chapter 4, in the context of delineation of the functional oligomerisation of the two proteins. Finally, in Chapter 5 an investigation of the DNA-binding properties of H-NS is presented.

All evidence presented in this thesis suggests a high structural similarity between the two proteins, and the biophysical techniques used did not identify any significant differences between H-NS and StpA. Nonetheless there are clear differences *in vivo* between the two proteins in terms of their regulation, and to a lesser degree, their function, suggesting that they possess different intracellular roles. Whilst the expression of H-NS is relatively insensitive to changes in environment, the opposite is true for StpA (see Sections 1.6 and 1.12). SEC experiments probing the interaction between H-NS and StpA (both via the coiled-coil interface and through the 'head-to-tail' mechanism of higher order oligomerisation) suggest that H-NS and StpA interact with greater affinity for each other (or at least as great) compared to the equivalent homomeric interaction. This suggests that *in vivo*, any free or newly synthesised StpA is rapidly incorporated into the highly abundant higher order oligomers of H-NS present in the cell. Any StpA that is not in complex with H-NS is then rapidly degraded by Lon protease (Johansson and Uhlin, 1999). The DNA-binding properties of these higher order heterooligomeric complexes of H-NS and StpA may then depend on the ratio of H-NS to StpA, as StpA displays different DNA- and RNA-binding properties compared to H-NS. Thus StpA may play a key role in the modulation of the global repressive effect of H-NS in response to changes in the environment by fine-tuning the DNA-binding properties of the large nucleoprotein complexes present in the nucleoid.

Chapter 2

Materials and methods

2.1 General chemicals and laboratory equipment

All chemicals were purchased from Sigma-Aldrich or BDH unless stated otherwise. All aqueous solutions were made using highly purified water from the ELGA Option 3 water purification system. DNA Miniprep kits and PCR product purification kits were purchased from Qiagen. Carbenicillin and isopropyl-B-D-thiogalactopyranoside (IPTG) were purchased from Melford Laboratories. All restriction enzymes were purchased from New England Biolabs. The *E. coli* XL1-Blue strain (Stratagene) and BL21 (DE3) pLysS strain (Stratagene) were used for all cloning and protein expression protocols, respectively. The pET14b bacterial expression vector (Novagen) was used in all cloning and protein expression protocols. All concentrations stated in this section are specified as weight-per-volume (w/v) unless stated otherwise.

Centrifugations steps outlined in this thesis were carried out using the ALC PK130R (for 15ml and 50ml Falcon centrifuge tubes), the Eppendorf 5415R (for microfuge tubes), the Hettich EBA12 (for microfuge tubes) or the Sorvall 5C5B. The SS34 and GS3 rotors were used (for 40 ml and 400 ml centrifugation tubes respectively) for centrifugation in the Sorvall RC5B centrifuge.

2.1.1 Growth media

Unless stated otherwise, Luria-Bertani (LB) broth was used as a growth medium for all protocols. LB broth was consisted of 10g/L tryptone (Gibco), 5g/L yeast extract (Gibco), and 5g/L NaCl. LB-agar plates consisted of LB broth fortified with 15g/L agar (Gibco), poured into Petri dishes.

Unless stated otherwise, all growth media for the *E. coli* XL1-Blue strain harbouring any pET14b plasmid was fortified with carbenicillin at 200 μg/ml concentration. All growth media for the *E. coli* BL21 (DE3) pLysS strain harbouring any pET14b plasmid was fortified with chloramphenicol and carbenicillin at 34μg/ml and 200μg/ml concentration respectively.

2.1.2 Minimal media

The use of minimal media allows selective uniform isotopic enrichment of the appropriate expressed protein with either ^{15}Nitrogen or ^{13}Carbon, or both. The per-litre quantities of the following chemicals are first dissolved in water: 6 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl and 0.7g NH_4Cl ($^{15}NH_4Cl$ if isotopic enrichment of the overexpressed protein in ^{15}N was required). This solution was then pH-adjusted to pH 7.0 and autoclaved. The following solutions were filter-sterilised through a 0.22μm bore filter before addition to the minimal media: 2 ml of 1 M $MgSO_4$, 10 μl 1 M $CaCl_2$, 1 ml 0.01 M $FeSO_4$ and 10 ml 20% glucose (or ^{13}C-glucose is isotopic enrichment of the overexpressed protein if ^{13}C was required).

2.2 General protocols for the manipulation of DNA

2.2.1 Electrophoresis of DNA

Agarose gel electrophoresis was carried out to characterise or purify DNA. Agarose was dissolved by heating in TAE buffer (40 mM Tris-Acetate, 1 mM EDTA), poured and cast into a gel. 1.5% agarose gels were used when separating DNA fragments smaller than 1 kilobase (kb) in size: otherwise 0.75% agarose gels were used to resolve the DNA. Appropriate quantities of DNA samples were mixed with 5x gel loading buffer (25% glycerol, 5 mM EDTA, 0.2% Bromophenol Blue and 0.2% Xylene Cyanol). The 100 bp DNA ladder (NEB) was used as a molecular weight standard. The buffered DNA samples and molecular weight standard were electrophoresed in TAE buffer at 100 mA. After electrophoresis the whole gel was stained in a

solution containing 0.2μg/ml of ethidium bromide for 20 minutes and destained by washing in water for 10 minutes. The DNA was then visualised by illuminating the gel with ultraviolet (UV) light.

2.2.2 DNA extraction from agarose gels

DNA excised from agarose gels was purified using the QIAquick gel extraction kit (Qiagen) as instructed in the product manual. This protocol was suitable for the purification of DNA fragments for use in DNA cloning experiments. In summary, the agarose gel slice containing the DNA fragment of interest is melted and passed through a silica gel membrane, to which the DNA selectively binds under the appropriate conditions. All contaminants are washed away, allowing subsequent elution of the purified DNA fragment in water.

2.2.3 Plasmid propagation and purification

Plasmid DNA was prepared using the Plasmid mini kit (Qiagen) as instructed in the product manual. Plasmid DNA was prepared from 5 ml of *E. coli* XL1-Blue cultures in LB media containing the appropriate plasmid, grown overnight with shaking. In summary, the addition of an alkaline buffer lyses the host cells, releasing the plasmid into solution. Much of the impurities precipitate at this stage and are removed by centrifugation. The sample is then passed through a silica-based gel, to which the plasmid DNA selectively binds under the appropriate conditions. Remaining impurities are removed by washing steps, allowing subsequent elution of the plasmid DNA in water. The purified plasmid DNA can then be used for sequencing, transformation and cloning.

2.3 Expression vectors for H-NS and StpA constructs

In addition to the full-length protein constructs of H-NS ($H\text{-}NS_{FL}$) and StpA ($StpA_{FL}$), several N-terminal and C-terminal constructs of H-NS and StpA were studies as described in this thesis. The two truncated H-NS constructs studied are: an N-terminal construct containing the first 64 residues ($H\text{-}NS_{1\text{-}64}$: the subscript denotes the equivalent residues on the full-length protein) and an N-terminal construct with the first 89 residues ($H\text{-}NS_{1\text{-}89}$). The two truncated StpA constructs studied are: an N-terminal construct containing the first 65 residues ($StpA_{1\text{-}65}$) and a C-terminal construct contain the last 44 residues ($StpA_{91\text{-}134}$).

2.3.1 H-NS_{FL} expression vector

A DNA insert corresponding to the *hns* gene from *S. typhimurium* (SWISS-PROT primary accession number P17428) was previously cloned into the expression vector pET14b (Novagen) between the 5' *Nde*1 site and the 3' *Bam*H1 site. The promoter region upstream of the cloned *hns* contains the T7 promoter sequence, the expression from which requires T7 polymerase. The pET14b vector contains the β-lactamase gene that confers resistance to the antibiotic carbenicillin. The use of the pET14b vector results in the expression of H-NS_{FL} as a fusion protein with an N-terminal His_6-tag. A thrombin cleavage site is located between the recombinant His-tag and the H-NS_{FL} protein allowing the partial removal of the affinity tag after expression and purification of the protein. This pET14b vector encoding HNS_{FL} was kindly provided by J. Hinton (Institute of Food Research).

2.3.2 H-NS_{1-64} and H-NS_{1-89} expression vectors

A DNA insert corresponding to the first 64 N-terminal residues of H-NS from *S. typhimurium* was previously cloned into the expression vector pET14b (Novagen) between the 5' *Nde*1 site and the 3' *Bam*H1 site. A pET14b expression vector encoding a polypeptide containing only the first N-terminal 89 residues of H-NS (H-NS_{1-89}) from *S. typhimurium* was similarly constructed. These pET14b vectors were kindly provided by C. Smyth.

2.3.3 $StpA_{FL}$ expression vector

A DNA insert corresponding to the *stpA* gene from *E. coli* (SWISS-PROT primary accession number P30017) was previously cloned in the expression vector pET14b (Novagen) between the 5' *Nde*1 site and the 3' *Bam*H1 site. This pET14b vector encoding $StpA_{FL}$ was kindly provided by A. Petrovic.

2.3.4 The construction of $StpA_{91-134}$ expression vector

Oligonucleotide primers were designed to bind regions flanking the coding DNA sequences corresponding to residues 91 to 134 of amino acid sequence of StpA. The restriction sites *Nde*1 and *Bam*H1 were introduced at the 5' and 3' ends of the coding sequence respectively. A stop codon (TAA) was inserted directly to the 3' end of the *Bam*H1 restriction site of the reverse primer. 9 'spacer' nucleotides were added to the 5' end of the restriction sites on the primers in order to ensure efficient digestion by

the relevant restriction enzyme. The polymerase chain reaction (PCR) primers were synthesised by MWG-Biotech. The oligonucleotide primers were designed such that the melting temperature (T_m) of the region annealing to the template was greater than 50°C, in order to minimise the probability of non-specific binding of oligonucleotide to template during PCR. The T_m was estimated using the formula, where A, T, G and C represent the bases adenine, thymine, guanine and cytosine respectively:

$$T_m\ (°C) = 2(A+T) + 4(G+C) \qquad \text{Equation 2.1}$$

$StpA_{91\text{-}134}$ forward primer
GGC CGT TAT CTA CAT ATG CAG CCG CGT CCG GCG
$StpA_{91\text{-}134}$ reverse primer
TAC TAG TAT GGA TCC TTA GAT CAG GAA ATC GTC G

A 100 µl PCR sample were set up containing 1x ThermoPol buffer (NEB), 0.5 µM of forward and reverse primer, 200 µM each of dNTP and 100 ng of plasmid DNA template encoding $StpA_{FL}$. The PCR amplification was carried out on a MWG-Biotech Primus thermocycler. The PCR sample was subjected to a 10 minute 'hot-start' where the reaction vessel is heated at 98°C, after which 2 units of $Vent_R$ DNA polymerase (NEB) is added to the reaction vessel. The sample was then subjected to 30 cycles of the following protocol: denaturation at 95°C for 30 seconds, annealing at 60°C for 45 seconds followed by extension at 72°C for 2 minutes. The resulting PCR products were purified by electrophoresis using 1.5% agarose gel, followed by extraction of the PCR amplification product using the QIAquick gel extraction kit (Qiagen).

The PCR amplification products were digested by the *Nde*1 and *Bam*H1 restriction enzymes (NEB). A 100 µl sample was set up containing 1x BamH1 buffer, 50 µl of purified PCR amplification product, BSA at 100 µg/ml concentration, and 40 units each of the restriction enzymes BamH1 and Nde1. A similar sample was prepared, containing pET14b plasmid at 100µg/ml concentration instead of PCR amplification product. The restriction digestion was carried out at 37°C for 2 hours. The digestion products were purified by agarose gel electrophoresis and extracted using the QIAquick gel extraction kit (Qiagen).

The PCR amplification products and the open plasmid were then ligated to form the pET14b expression vector encoding $StpA_{91-134}$. An 18 µl sample was prepared, containing 5 µl of purified restriction digested DNA insert corresponding to $StpA_{91-134}$ and 0.5 µl of purified restriction digested pET14b vector. This mixture was heated at 45°C for 10 minutes and then cooled on ice. 2 µl of 10x ligase buffer and 200 units of T4 ligase (NEB) were added to the ligation reaction mixture. The ligation mixture was incubated at 16°C for 15 hours. The ligation reaction was terminated by heating the mixture at 65°C for 30 minutes.

2.3.5 Construction of the $StpA_{1-65}$ expression vector

A stop codon was introduced into the 66th position of the amino acid sequence of StpA by site-directed mutagenesis, thereby creating an N-terminal construct containing the first 65 amino acids. The Quikchange site-directed mutagenesis kit (Stratagene) was used to introduce the mutation, as detailed in the product manual. In summary, two oligonucleotide primers, each complimentary to the opposite strand of the plasmid DNA template, are designed so that they contain the desired mutation. Extension from these oligonucleotide primers by the *PfuTurbo* DNA polymerase during the course of the temperature cycling results in non-methylated, mutated plasmid. The reaction mixture is then treated with *Dpn*1 endonuclease which is specific for the methylated parental plasmid DNA, leaving only the mutated plasmid DNA. This plasmid is then transformed into *E. coli* XL1-Blue strain.

$StpA_{1-65}$ forward primer
GGC TGG AGC TGA TGA AAT GAG ACG GAA TTA ACC CGG
$StpA_{1-65}$ reverse primer
CCG GGT TAA TTC CGT CTC ATT TCA TCA GCT CCA GCC

2.4 Producing competent cells

Competent cells of the XL1-Blue strain of *E. coli* were prepared using the following protocol. Cells from glycerol stocks of *E. coli* XL1-Blue strain were streaked onto an LB-agar plate containing carbenicillin (200 µg/ml) and incubated overnight at 37°C. 100 ml of LB medium (containing 200 µg/ml carbenicillin) was inoculated with a single colony of *E. coli* XL1-Blue strain and grown overnight at 37°C with shaking at 200 rpm. The cells were pelleted with centrifugation (2000 g, 15 minutes, 4°C) and

resuspended in ice-cold, sterile 0.1M $CaCl_2$ solution with gentle agitation. This suspension of *E. coli* XL1-Blue strain cells were then left on ice for 20 minutes. The cells were pelleted by centrifugation (750 g, 10 minutes, and 4°C), resuspended in ice-cold sterile $CaCl_2$ and left on ice for a further 20 minutes. This suspension was pelleted by centrifugation (750 g, 10 minutes, 4°C) and resuspended into 4 ml of sterile ice-cold 0.1M $CaCl_2$, 10% (v/v) glycerol. These competent cells were divided onto 100 µl aliquots in 1.5 ml microfuge tubes and stored at –80°C.

Competent *E. coli* BL21 (DE3) pLysS strain cells were prepared as outlined above, with the exception of the supplementing of the LB medium and LB-agar plates with chloramphenicol (at 34 µg/ml) instead of carbenicillin.

2.5 Transformation

1 µl of appropriate plasmid DNA (at 100 µg/ml concentration) was added to 100 µl of the gently thawed competent cells (either the XL1-Blue strain or the BL21 (DE3) pLysS strain) contained in a 1.5 ml microfuge tube. The suspension of cells and plasmid DNA in the tube were left on ice for 30 minutes, heat-shocked by incubation at 42° for 45 seconds, and then returned to ice for 2 minutes. 900 µl of LB media pre-warmed to 42°C was added to the suspension of cells. The tubes were incubated at 37°C for 30 minutes with shaking at 200 rpm. The cells within the microfuge tubes were pelleted with centrifugation at 16,000g for 1 minute and resuspended in 100 µl of LB medium. The cells were spread over LB-agar plates containing the appropriate antibiotics. These plates were incubated at 37°C overnight.

2.6 Large scale protein expression

This protocol for the overexpression of protein applies to all constructs of H-NS and StpA described in this thesis. Where unlabelled protein was required, LB broth was used as a growth medium. Where either ^{15}N or ^{13}C isotopic enrichment (or enrichment in both isotopes) was required, the appropriate minimal medium was used (see Section 2.1.2).

A colony of *E. coli* BL21 (DE3) pLysS strain transformed with the appropriate pET14b plasmid was picked from a LB-agar plate and used to inoculate 10 ml of growth medium (either LB or minimal medium) containing carbenicillin and

chloramphenicol at concentrations 200 μg/ml and 34 μg/ml, respectively. This growth medium was incubated at 37°C with shaking at 200 rpm for 4 to 5 hours until the optical density at 600 nm (OD_{600}) was 0.5. 50 μl of the culture was then used to inoculate 100 ml of fresh growth medium (either LB or minimal medium) containing carbenicillin and chloramphenicol at 200 μg/ml and 34 μg/ml, respectively. This culture was shaken overnight at 37°C at 200 rpm. 10 ml aliquots of this overnight culture was used to inoculate 2-litre flasks each containing 500 ml of fresh growth medium (either LB or minimal medium) supplemented with carbenicillin and chloramphenicol at 200 μg/ml and 34 μg/ml, respectively. The flasks were incubated at 37°C shaken at 200 rpm until the OD_{600} reached 0.5. Protein expression was then induced with 1 mM IPTG. The flasks were incubated for a further 4 hours at 37°C shaken at 200 rpm.

The growth media was centrifuged at 3500 g at 4°C for 30 minutes in a GS-3 rotor in a Sorvall RC5B centrifuge. The supernatant was discarded. The cell pellets were re-suspended in 40 ml of lysis buffer (20 mM Tris-HCl, 500 mM NaCl, pH 7, 1% Triton).

2.7 Large scale purification

The purification protocols depended on the protein construct, as summarised in table 2.1. All purifications were carried out using an $AKTA_{FPLC}$ protein purification system (Amersham biosciences). The $AKTA_{FPLC}$ system allowed detection of protein by absorbance at 280 nm. All dialysis steps were carried out using dialysis membrane

	First purification step	Removal of His-tag	Second purification step
$H\text{-}NS_{FL}$	Metal affinity chromatography	Yes	SP sepharose chromatography
$H\text{-}NS_{1\text{-}64}$	"	"	Preparative SEC
$H\text{-}NS_{1\text{-}89}$	"	"	-
$StpA_{FL}$	"	No	DNA-cellulose chromatography
$StpA_{1\text{-}65}$	"	Yes	Preparative SEC
$StpA_{91\text{-}134}$*	"	"	SP sepharose chromatography

Table 2.1
Summary of the purification protocols of the various constructs of H-NS and StpA.
* Valuable samples such as [^{13}C, ^{15}N]-$StpA_{91\text{-}134}$ and [^{15}N]-$StpA_{91\text{-}134}$ were subjected to a third purification step of preparative SEC.

with a MW cut-off of 3.5 kDa (Spectrapore). All chromatography resins were poured into XK16/20 columns (Amersham biosciences) unless stated otherwise. Purification steps were generally closely monitored by checking fractions with SDS-PAGE. The concentrations of all protein constructs used in this study were increased using Centriprep YM-3 centrifugal concentrators, with a 3 kDa MW cut-off (Amicon), by centrifugation at 2000 g at 4 °C. Examples of typical expression and purification procedures are shown in Figures 2.1, 2.2 and 2.3, where various samples during the purification of $StpA_{FL}$, $StpA_{91-134}$, and $StpA_{1-65}$ respectively are analysed by SDS-PAGE.

2.7.1 Affinity chromatography of His-tagged proteins

As all H-NS and StpA constructs described in this thesis have a recombinant N-terminal His-tag, the initial purification steps encompassing cell lysis and metal affinity purification described in this sub-section are common to all constructs.

Immediately prior to cell lysis one tablet of the 'complete, EDTA-free protease inhibitor cocktail tablets' (Roche) was added to every 25 ml of lysis buffer. The *E. coli* BL21 (DE3) pLysS strain cells containing the overexpressed protein were lysed by sonication using 40 cycles of 10 second sonication pulses followed by a rest period of 20 seconds, using a Soniprep 150 (Sanyo MSE). The cell pellets were kept in ice during the duration of sonication. The resultant cell lysate was cleared of insoluble cell debris by centrifugation at 25,000 g in a SS34 rotor at 4°C for 1 hour in a Sorvall RC5B centrifuge.

TALON metal affinity resin (Clontech), packed into a XK16/20 column (Amersham Pharmacia) and attached to an $AKTA_{FPLC}$ protein purification system (Amersham Pharmacia), was equilibrated by flowing buffer through the column (20 mM Tris-HCl pH 8, 500 mM NaCl). Protein elution was detected by absorbance at 280 nm. Cell lysate was loaded onto the TALON resin at 1 ml/min. The flow-through fraction, shown not to contain any recombinant protein by analysis by SDS-PAGE, was discarded. The column was washed with a further 10 column volumes of buffer. Non-specifically binding proteins were washed off the column with 10 column volumes of wash buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, and 20 mM imidazole). The His-tagged protein was eluted in 2 column volumes of elution buffer (20 mM Tris-HCl pH

8, 500 mM NaCl, and 250 mM imidazole). The His-tagged protein solution was sealed inside dialysis membrane (3.5 kDa MW cut-off, Spectrapore) and buffer-exchanged with 20 mM sodium phosphate pH 7, 500 mM NaCl, and 1 mM EDTA.

2.7.1.1 Removal of the His-tag

In this sub-section is described a protocol for the proteolytic removal of the N-terminal His-tag from overexpressed protein. This protocol is applicable to all H-NS and StpA constructs, except $StpA_{FL}$, which without the His-tag, was found to precipitate rapidly under all conditions tested.

The His-tagged protein was sealed inside a dialysis membrane (3.5 kDa MW cut-off, Spectrapore) and buffer-exchanged with thrombin cleavage buffer (20 mM Tris-HCl pH 8, 300 mM NaCl). High activity bovine thrombin (Calbiochem, 1 unit of enzyme per 2 mg of His-tagged protein) was added to the protein solution. After incubation for 3 hours at room temperature the cleavage reaction was terminated by the addition of 1 tablet of 'complete, EDTA-free protease inhibitor cocktail tablets' (Roche) to every 50 ml of thrombin cleavage buffer. The protein solution was passed through the TALON column once more to remove any remaining His-tagged species.

2.7.2 Cation exchange chromatography

The protocol involving purification of protein using cation exchange described in this sub-section was applied to $StpA_{91-134}$ and $H\text{-}NS_{FL}$. The solution of overexpressed protein was diluted with running buffer (20 mM Tris-HCl, 10 mM NaCl, pH 8) in such a way that the total concentration of NaCl was less than 50 mM. This solution was loaded onto a column packed with fast-flow SP sepharose beads (Amersham Pharmacia) pre-equilibrated with buffer (20 mM Tris-HCl, 10 mM NaCl, pH 8). After washing the beads with buffer the bound protein is eluted with 20 mM Tris-HCl, 1M NaCl, pH 8.

2.7.3 Preparative size exclusion chromatography

The protocol described in this sub-section was applied during the purification processes of $StpA_{1-65}$, $StpA_{91-134}$ and $H\text{-}NS_{1-64}$. Preparative SEC was carried out using a Sephadex 75 column (XK 16/60, Amersham Pharmacia) pre-equilibrated with buffer (20 mM Tris-HCl pH 8, 500 mM NaCl). The concentration of the protein loaded was

manipulated such that the maximum volume of samples loaded was 5 ml. The column was run at a flow rate of 1.0 ml/min and 2 ml fractions were collected. Protein elution was detected using UV absorbance at 280 nm.

SEC of H-$NS_{1\text{-}89}$ was carried out at different temperatures using the Sephadex 75 column. 2 ml samples containing 168 μM H-$NS_{1\text{-}89}$, 20 mM sodium phosphate pH 7.0 and 300 mM NaCl were injected into the column with a flow rate of 1.5 ml/min. Protein elution was detected by absorbance at 280 nm. The column, equipped with a thermostatic jacket, was kept at a constant temperature (± 0.1°C) using an external water bath. Figure 2.4 shows a plot of the logarithm of the molecular weight against the elution volume of albumin (66 kDa), carbonic anhydrase (29 kDa), cytochrome C (12.4 kDa) and aprotinin (6.5 kDa).

2.7.4 DNA-cellulose chromatography

In this sub-section a protocol for the purification of $StpA_{FL}$ is described. A solution containing $StpA_{FL}$ was loaded onto a DNA-cellulose column pre-equilibrated with buffer (10 mM sodium phosphate, 10 mM Tris-HCl, pH 7.0, 500 mM NaCl). After loading, the column was washed with at least 5 column volumes of buffer. $StpA_{FL}$ was eluted using elution buffer (10 mM sodium phosphate, 10 mM Tris-HCl, pH 7.0, 1 M NaCl). Protein elution was detected using UV absorbance at 280 nm.

2.8 Measurement of protein concentration by UV Spectroscopy

Concentrations of all constructs used in this study were determined using the Cary 50 UV-visible spectrophotometer (Varian), in quartz cuvettes of 1 cm pathlength (Hellma). Molar extinction coefficients were estimated on the basis of the amino acid sequences of the various constructs (Gill and von Hippel, 1989), and are shown in table 2.2. Protein concentrations were estimated using the Beer-Lambert law,

$$A_{280} = \varepsilon_{280} \cdot c \cdot l \qquad \text{Equation 2.2}$$

where A_{280} is the absorbance of the protein sample at 280 nm, ε_{280} is the molar extinction coefficient at 280 nm ($mole^{-1}$ cm^{-1}) and l is the pathlength of the sample (cm).

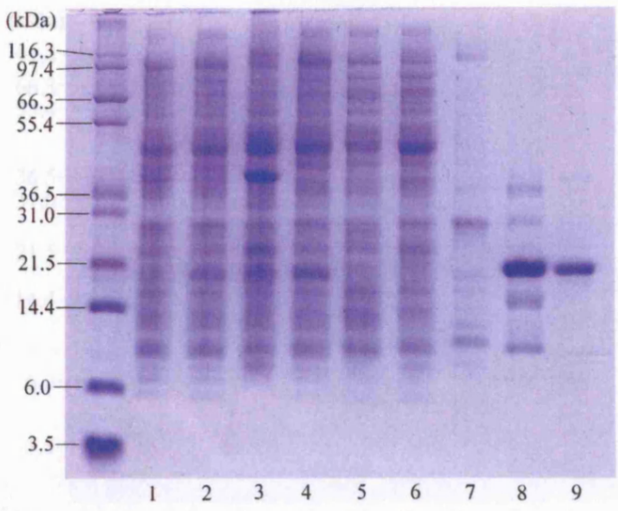


Figure 2.1
SDS-PAGE analysis of the purification of $StpA_{FL}$. The following samples are displayed: lane 1, whole cell lysate, before induction; lane 2, whole cell lysate, after induction; lane 3, insoluble cell debris after sonication and centrifugation; lane 4, soluble fraction after sonication and centrifugation; lane 5, metal-affinity column flow-through; lane 6, metal-affinity column flow-through; lane 7, wash with 20 mM imidazole buffer; lane 8, elution with 500 mM imidazole; lane 9, after purification with DNA-cellulose chromatography. The samples were electrophoresed in 12% SDS-PAGE gel, stained with coomassie blue.

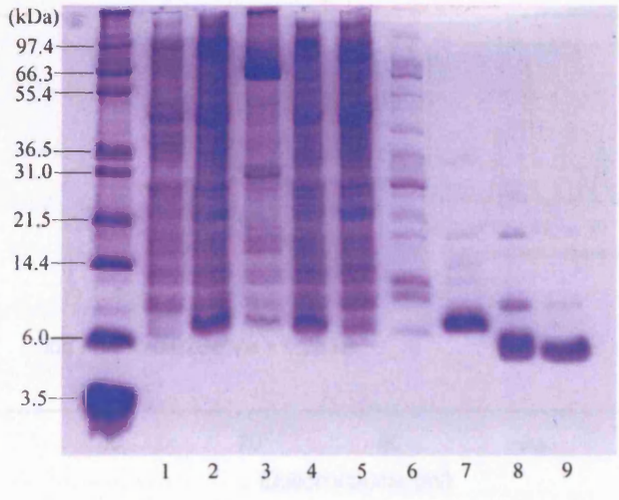


Figure 2.2
SDS-PAGE analysis of the purification of $StpA_{91-134}$. The following samples are displayed: lane 1, whole cell lysate, before induction; lane 2, whole cell lysate, after induction; lane 3, insoluble cell debris after sonication and centrifugation; lane 4, soluble fraction after sonication and centrifugation; lane 5, metal-affinity column flow-through; lane 6, wash with 20 mM imidazole buffer; lane 7, elution with 500 mM imidazole; lane 8, after thrombin cleavage; lane 9, after purification with SP sepharose. The samples were electrophoresed in 12% SDS-PAGE gel, stained with coomassie blue.

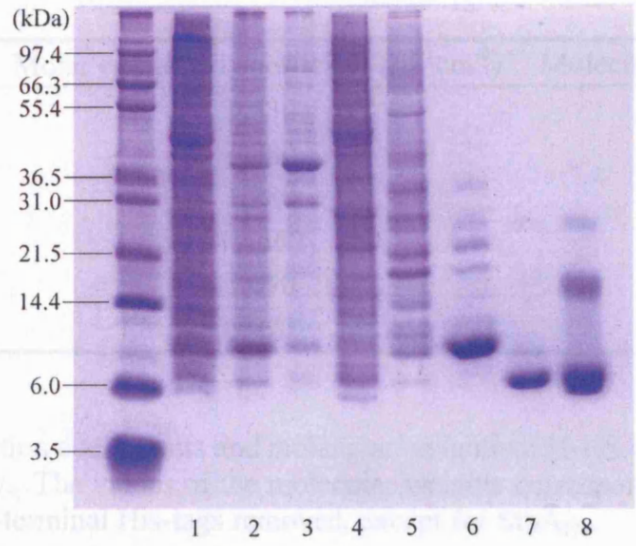


Figure 2.3
SDS-PAGE analysis of the purification of $StpA_{1-65}$. The following samples are displayed: lane 1, whole cell lysate, before induction; lane 2, whole cell lysate, after induction; lane 3, insoluble cell debris after sonication and centrifugation; lane 4, soluble fraction after sonication and centrifugation; lane 5, metal-affinity column flow-through; lane 6, elution with 500 mM imidazole buffer; lane 7, after purification with preparative SEC; lane 8, after thrombin cleavage. The samples were electrophoresed in 12% SDS-PAGE gel, stained with coomassie blue.

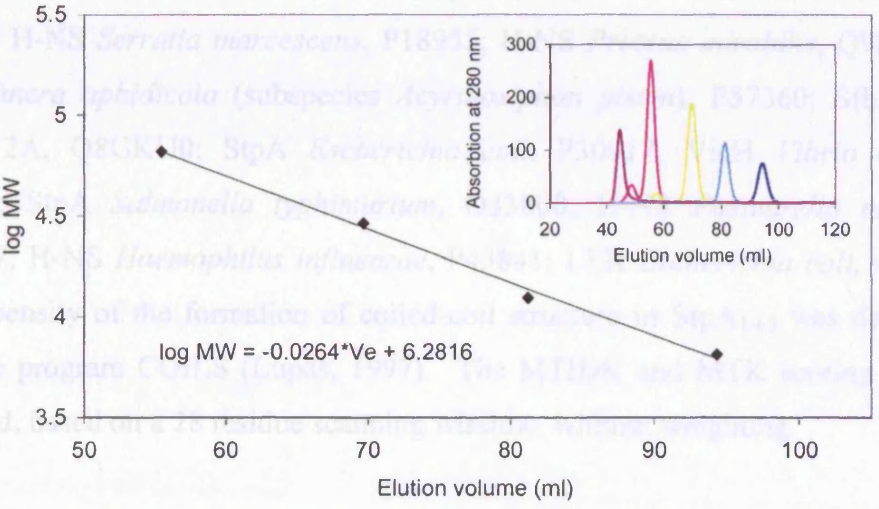


Figure 2.4
The logarithm of the molecular weight of globular protein standards plotted against elution volume using the Sephadex 75 column. Albumin (—, 66 kDa), carbonic anhydrase (—, 29 kDa), cytochrome C (—, 12.4 kDa) and aprotinin (—, 6.5 kDa) were used as standards. Blue dextran (—, ~1.5 MDa) was used to determine the void volume of the column. The data was fitted with linear regression, and the equation is shown above. The apparent molecular weight of a sample protein can be estimated by comparing the elution volume with the equation shown above.

	Molar extinction coefficient ($M^{-1}cm^{-1}$)	Molecular weight (Da)
$H\text{-}NS_{FL}$	9,530	15,824
$H\text{-}NS_{1\text{-}64}$	1,280	7,940
$H\text{-}NS_{1\text{-}89}$	1,280	10,528
$StpA_{FL}$	12,660	17,511
$StpA_{1\text{-}65}$	5,690	8,367
$StpA_{91\text{-}134}$	6,970	5,288

Table 2.2
The molar extinction coefficients and molecular weights of H-NS and StpA constructs used in this study. The values of the molecular weights correspond to constructs that have had their N-terminal His-tags removed, except for $StpA_{FL}$.

2.9 *In silico* sequence analysis

A selection of the members of the H-NS-like family were identified using BLAST 2.0 (Altschul *et al.*, 1997). Multiple sequence alignment of these proteins was carried out using ClustalW (Thompson *et al.*, 1994). The multiple alignment was graphically enhanced using Bioedit (Hall, 2004). The following proteins (with their corresponding accession numbers) are presented in the multiple alignment: H-NS *Salmonella typhimurium*, P17428; H-NS *Escherichia coli*, P08936; H-NS *Erwinia chrysanthemi*, O53007; H-NS *Serratia marcescens*, P18955; H-NS *Proteus mirabilis*, Q9L657; H-NS *Buchnera aphidicola* (subspecies *Acyrthosiphon pisum*), P57360; Sfh *Shigella flexneri* 2A, Q8GKU0; StpA *Escherichia coli*, P30017; VicH *Vibrio cholerae*, Q9Z4R1; StpA *Salmonella typhimurium*, O33800; H-NS *Pasteurella multocida*, Q9CMF7; H-NS *Haemophilus influenzae*, P43841; LER *Escherichia coli*, Q8VTF2. The propensity of the formation of coiled-coil structure in $StpA_{1\text{-}65}$ was determined using the program COILS (Lupas, 1997). The MTIDK and MTK scoring matrices were used, based on a 28 residue scanning window, without weighting.

2.10 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

H-NS and StpA samples were routinely examined using Tris-tricine SDS-PAGE, in the context of monitoring purification procedures and the analysis of the His-tag pull-down assay samples. SDS-polyacrylamide gels consisting of a 4% stacking layer and 12% resolving layer were used throughout the study. The constituents of the gel are shown in Table 2.3. Electrophoreses was carried out on the Mini-Protean II system

(Bio-Rad) linked to a PowerPac 300 power supply (Bio-Rad). Protein samples were mixed with 5x SDS sample buffer (0.5 M Tris-HCl pH 6.8, 10% SDS, 0.5% bromophenol blue, 50% glycerol), heated at 100°C for 5 minutes, and centrifuged at 16,000 g (to remove insoluble matter) for 10 minutes. 10 µl of buffered sample was loaded onto the gel and electrophoresed (at 60 mA per gel) for approximately 80 minutes. Mark 12 unstained standards (Invitrogen) were concurrently electrophoresed with protein samples as a molecular weight standards. After completion of electrophoresis the gels were stained in a solution containing 0.1% coomassie brilliant blue R, 40% methanol (v/v) and 10% acetic acid (v/v) for 1 hour, followed by 3~4 successive 100 ml washes with destaining solution containing 40% methanol (v/v) and 10% acetic acid (v/v).

	12% resolving gel	4% stacking gel
Gel buffer (3M Tris-HCl pH 8.45, 0.3% SDS)	10 ml	4 ml
30% acrylamide solution (37.5: 1 ratio of acrylamide: bisacrylamide)	11.25 ml	1.6 ml
Glycerol	3.35 ml	-
10% ammonium persulphate	90 µl	50 µl
N, N, N', N'-Tetramethylethylenediamine (TEMED)	25 µl	20 µl
Water	3.35 ml	6.4 ml

Table 2.3
Recipe of 12% Tris-tricine SDS-PAGE gels.

2.11 Analytical SEC

Analytical SEC experiments were carried out on a Superose 12 HR 10/30 (Amersham Biosciences) attached to an $AKTA_{FPLC}$ protein purification system (Amersham Biosciences). Samples were injected onto the pre-equilibrated column in volumes depending on the application, and varied between 100 µl to 1 ml. The flow-rate was set at 0.4 ml/min, and protein elution was detected by absorbance at 280 nm. Experimental conditions such as buffer conditions and protein concentrations differed between experiments and are presented alongside the results.

The column was calibrated using the gel filtration standard kit (Bio-Rad), containing markers with molecular weights ranging from of 1.35 kDa to 158 kDa, dissolved in 20 mM sodium phosphate buffer pH 7.0, 500 mM NaCl. Figure 2.5 shows the relationship between the logarithms of molecular weights of the various standards plotted against elution volume.

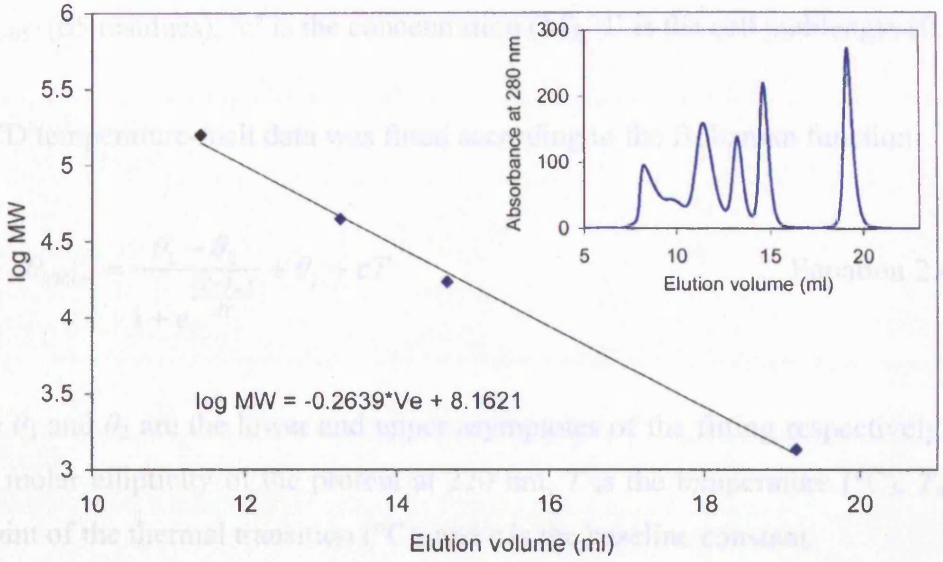


Figure 2.5
The logarithm of the molecular weight of globular protein standards plotted against elution volume using the Superose 12 column. Bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), equine myoglobulin (17 kDa) and vitamin B_{12} (1.35 kDa) were used as standards. Thyroglobulin (670 kDa) was used to determine the void volume of the column. The data was fitted with linear regression, and the equation is shown above. The apparent molecular weight of a sample protein can be estimated by comparing the elution volume with the equation shown above.

2.12 Circular dichroism spectroscopy

Circular dichroism spectroscopy (CD) was carried out on 30 μM $StpA_{1-65}$ samples in 20 mM sodium phosphate pH 7.0 buffer in a variety of concentrations of NaCl using an Aviv CD spectrophotometer, model 202SF. CD spectra were recorded between 180 nm and 260 nm in 1 nm increments with an averaging time of 10 seconds. The absorbance of the $StpA_{1-65}$ sample was concurrently measured; data with a corresponding dynode voltage greater than 600 volts was discarded. CD spectra were acquired at 10 °C increments between 5 °C and 95 °C. Buffer scans were subtracted from the corresponding $StpA_{1-65}$ CD scans. CD spectra were recorded in 0.1 cm path

length quartz cuvettes (Hellma). CD data was represented as per-residue molar ellipticity:

$$\text{Per-residue molar ellipticity} = \theta_R (\#\ \text{aa} \cdot \text{c} \cdot \text{l} \cdot 10)^{-1} \qquad \text{Equation 2.3}$$

where 'θ_R' is the raw CD data (millidegrees), '# aa' is the number of amino acids in $StpA_{1\text{-}65}$ (65 residues), 'c' is the concentration (M), 'l' is the cell pathlength (0.1 cm).

The CD temperature-melt data was fitted according to the Boltzman function:

$$\theta_{220nm} = \frac{\theta_1 - \theta_2}{1 + e^{\frac{(T - T_m)}{dT}}} + \theta_2 + cT \qquad \text{Equation 2.4}$$

where θ_1 and θ_2 are the lower and upper asymptotes of the fitting respectively, θ_{220nm} is the molar ellipticity of the protein at 220 nm, T is the temperature (°C), T_m is the midpoint of the thermal transition (°C), and c is the baseline constant.

2.13 Nuclear Magnetic Resonance Spectroscopy

All NMR spectra were recorded on the 500 MHz and 600 MHz Varian spectrometers at 25 °C at the UCL/Ludwig Institute of Cancer Research NMR laboratory. All NMR experiments were set up with assistance from Dr M. Williams, Dr. R. Harris or J. Taylor. The pulse sequences used in NMR studies on $StpA_{91-134}$ are listed in Table 2.4, along with the number of data points, sweep-widths and transmitter offsets used in each experiment.

[^{1}H, ^{15}N]-HSQC (Kay *et al.*, 1992a)		
Dimension	^{1}H	^{15}N
Number of points	1024	64
Sweep-width (Hz)	4000	1550
Transmitter offset (ppm)	6.440	120.335

HNCACB (Kay *et al.*, 1992a; Yamazaki *et al.*, 1994)			
Dimension	^{1}H	^{13}C	^{15}N
Number of points	512	72	32
Sweep-width (Hz)	4000	7650	1550
Transmitter offset (ppm)	6.440	44.827	120.335

CBCACONH (Kay and Muhandiram, 1994)			
Dimension	^{1}H	^{13}C	^{15}N
Number of points	512	46	32
Sweep-width (Hz)	4000	7650	1550
Transmitter offset (ppm)	6.440	44.820	120.335

CCONH (Grzesiek and Bax, 1993)			
Dimension	^{1}H	^{13}C	^{15}N
Number of points	512	64	26
Sweep-width (Hz)	4000	8050	1550
Transmitter offset (ppm)	6.440	37.866	120.335

Table 2.4
NMR pulse sequences used in NMR studies of $StpA_{91-134}$.

HCCONH (Grzesiek and Bax, 1993)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{15}N
Number of points	512	90	32
Sweep-width (Hz)	4000	6000	1550
Transmitter offset (ppm)	6.440	4.750	119.842

HNHA (Kuboniwa *et al.*, 1994; Vuister and Bax, 1993)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{15}N
Number of points	1024	40	24
Sweep-width (Hz)	6000	4350	1750
Transmitter offset (ppm)	4.790	6.240	119.842

HNHB (Archer *et al.*, 1991; Bax *et al.*, 1994)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{15}N
Number of points	1024	48	32
Sweep-width (Hz)	6000	6000	1750
Transmitter offset (ppm)	4.790	4.790	119.842

[^{1}H, ^{13}C]-CT-HSQC (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991; Vuister and Bax, 1992)		
Dimension	^{1}H	^{13}C
Number of points	1024	256
Sweep-width (Hz)	5900	11000
Transmitter offset (ppm)	4.790	45.822

HCCH-TOCSY (Kay *et al.*, 1993)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{13}C
Number of points	1024	64	72
Sweep-width (Hz)	5200	5200	8050
Transmitter offset (ppm)	3.379	3.379	43.037

Table 2.4, continued
NMR pulse sequences used in NMR studies of $StpA_{91-134}$.

[^{1}H, ^{13}C]-CT-HSQC, for aromatic side chains (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991; Vuister and Bax, 1992)		
Dimension	^{1}H	^{13}C
Number of points	1024	64
Sweep-width (Hz)	5200	2500
Transmitter offset (ppm)	3.379	129.601

HCCH-TOCSY, for aromatic side chains (Kay *et al.*, 1993)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{13}C
Number of points	1024	48	40
Sweep-width (Hz)	5200	2500	5200
Transmitter offset (ppm)	3.379	4.750	129.601

^{13}C-NOESY-HSQC, for aromatic side chains (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{13}C
Number of points	1024	72	36
Sweep-width (Hz)	5200	6000	2500
Transmitter offset (ppm)	3.379	4.750	129.601

CBHD (Yamazaki *et al.*, 1993)		
Dimension	^{1}H	^{13}C
Number of points	1024	38
Sweep-width (Hz)	6000	4800
Transmitter offset (ppm)	4.740	38.900

^{15}N-T_1 (Kay *et al.*, 1989; Kay *et al.*, 1992b; Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)		
Dimension	^{1}H	^{15}N
Number of points	1024	64
Sweep-width (Hz)	4000	1600
Transmitter offset (ppm)	6.739	119.330

Table 2.4, continued
NMR pulse sequences used in NMR studies of $StpA_{91-134}$.

^{15}N-T_2 (Kay *et al.*, 1989; Kay *et al.*, 1992b; Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)		
Dimension	^{1}H	^{15}N
Number of points	1024	64
Sweep-width (Hz)	4000	1600
Transmitter offset (ppm)	6.449	119.330

[^{1}H, ^{15}N]-HNOE (Barbato *et al.*, 1992; Kay *et al.*, 1989; Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)		
Dimension	^{1}H	^{15}N
Number of points	1024	128
Sweep-width (Hz)	4000	1600
Transmitter offset (ppm)	6.790	119.330

^{15}N-TOWNY-HSQC (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{15}N
Number of points	1024	128	32
Sweep-width (Hz)	4000	6000	1600
Transmitter offset (ppm)	6.440	4.750	119.05

^{15}N-NOESY-HSQC (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{15}N
Number of points	1024	128	32
Sweep-width (Hz)	4100	6000	1750
Transmitter offset (ppm)	4.750	4.750	119.842

^{13}C-NOESY-HSQC (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991)			
Dimension	Direct ^{1}H	Indirect ^{1}H	^{13}C
Number of points	1024	90	64
Sweep-width (Hz)	8000	7000	9900
Transmitter offset (ppm)	2.88	4.790	43.037

Table 2.4, continued
NMR pulse sequences used in NMR studies of $StpA_{91-134}$.

2.13.1 NMR data processing with NMRPipe

All NMR spectra were first referenced to 2,2-dimethyl-2-silapentane-5-sulphonate (DSS), using an in-house program called 'newoff', written by Dr. Mark Pfuhl. All NMR data were then processed using the NMRPipe program (Delaglio *et al.*, 1995). NMRPipe was used to apply various filters to the datasets, for example, first and second order phase corrections, window functions to optimise spectral resolution, zero-filling of datasets, linear prediction of data in the indirect dimension, suppression of the water signal, Fourier transformation of the datasets and baseline corrections. The various methods used to process the data are described in the following sections. The precise schemes used varied between different NMR experiments. Finally, NMRPipe allowed conversion of processed data into formats readable by NMR spectral analysis programs, including plot2 (Boucher, 2002), nmrDraw (Delaglio *et al.*, 1995) and ANSIG (Kraulis, 1989; Kraulis, 1994). The peak maxima of the NMR spectra were identified using an in-house program called 'mp_peakpick.al', written by Dr. Mark Pfuhl. Peak intensities of the NOESY-HSQC spectra were calculated using an in-house program called 'mw_height_spectrum.al', written by Dr. Mark Williams.

2.13.1.1 Linear prediction

The analogue NMR signal from an experiment must be digitised during acquisition before storage and manipulation of the data by computers; in other words, the continuous analogue signal is sampled at defined intervals to give a digitised signal. Due to time limits in recording single experiments, especially in 3-dimensional (3D) NMR experiments, the resolution of the data must be sacrificed. The linear prediction function in NMRPipe allows improvement of the resolution in the indirect dimensions by predicting data points based on existing ones. The linear prediction function was used to double the number of data points in the indirect dimensions of all NMR spectra of $StpA_{91-134}$.

2.13.1.2 Apodisation

One use of apodisation is to counter the artefacts of the Fourier transformation algorithms. Often NMR experiments are recorded with limited signal acquisition time, and can result in insufficient time for the FID to decay to zero. The rest of the FID is discarded, causing a truncation of exponential decay profile of the FID that amounts to the creation of a sinc function (i.e. a step function at the end of the FID). Subse-

quent Fourier transformation of the truncated FID results in distortion of the data, seen as a 'ringing' effect where the baseline near a cross-peak oscillates. Apodisation functions (also known as window functions), such as the Lorentzian, Gaussian and Sine bell functions, can be applied to the data to remove the step function at the end of the truncated FID, removing the distortion in the Fourier transform of the data, at the expense of lowering the resolution.

Apodisation can also improve NMR signals resulting from unfavourable relaxation; for example, when studying large proteins by NMR, signals can decay very quickly due to fast transverse relaxation. Fourier transforms of quickly relaxing signals give broad line-widths (see section 3.1.3). Apodisation can be used to modify the FID, so that the FID decays with an apparently longer transverse relaxation rate, resulting in sharper cross-peaks, at the expense of increasing the noise. Conversely, signal-to-noise ratios of the NMR data may be improved at the expense of resolution.

2.13.1.3 Zero-filling

The digital spectral resolution increases with longer acquisition times, in other words, a longer FID. A number of zeroes can be added onto the end of the FID in a process called zero-filling, thereby increasing the apparent FID acquisition length, without actually increasing the information content of that FID, resulting in the increase of the digital resolution of the dataset. Zero-filling also allows the imaginary part of the data to be transformed into the frequency domain, provided the length of the zero-filling is equal to or greater than the length of the original data.

After smoothing any truncation at the end of the FID by using the appropriate apodisation function (see Section 2.13.1.2), a number of zeroes equal to the number of data points in the FID were added to the end of the FID before Fourier transformation.

2.13.1.4 Other processing function of NMRPipe

NMRPipe enables many other signal processing functions, some of which are briefly described in this section.

Fast Fourier transformation, an algorithm based on Fourier transformation, allows interconversion of the NMR data between the FID (NMR data described in the time domain) and the spectrum (NMR data described in the frequency domain).

Zero order phase correction allows mixing of the phases of the real and imaginary part of the FID, in order to achieve a pure absorption real Lorentzian curve and a pure dispersion imaginary Lorentzian curve. The phase of the spectra must be corrected to allow accurate determination of chemical shifts and cross-peak intensities which are vital in the later stages of $StpA_{91-134}$ structure determination.

First order phase correction allows the correction of offset-dependant phase errors. Due to small imperfections in the NMR spectrometer hardware, phase errors in the recorded FID may occur, where the magnitude of the error depends on the frequency of the signal.

Polynomial baseline correction subtracts a polynomial baseline fitted to spectra in the frequency domain. Only zero order or first order baseline subtractions were applied to $StpA_{91-134}$ spectra.

2.13.2 Backbone Dihedral Prediction using TALOS

The TALOS program compares the chemical shift values of Hα, Cα, CB, CO and N of three consecutive residues and the residue type of the query protein with a database containing similar information derived from 20 proteins of known structure. If a significant number of triplets of residues of the same sequence with similar chemical shift values is found from the database to have consistent values of phi (φ) and psi (ψ) backbone torsion angles for the central residue of the triplet, then the TALOS program can provide a prediction of the unknown phi and psi angles of the query protein, as well as providing a degree of confidence of that prediction (Cornilescu *et al.*, 1999).

The amino acid sequence and Hα, Cα, Cβ, Co and N chemical shift values of $StpA_{91-134}$ are used as inputs to the TALOS program. The N-terminal and C-terminal residues are excluded from the 48 total amino acids of $StpA_{91-134}$ in this analysis, giving rise to 46 triplets. The amino acid sequence and chemical shift values of the Hα, Cα,

Cβ, CO and N nuclei of all three residues within the triplet are compared to a database (giving rise to 18 independent variables: three amino acid identities and 15 chemical shift values). The phi and psi backbone dihedral angles of the central residues of the ten best matching triplets are then plotted on a Ramachandaran diagram. If all ten pairs of backbone dihedral angles of the best database matches lie within the same region of the Ramachandaran diagram, then a high level of confidence is assigned to the matches. The mean and standard deviation of the ten pairs of phi and psi backbone dihedral angles were calculated and used as restraints in subsequent structure calculations. If only nine (or less) pairs of phi and psi backbone dihedral angles of the ten best database matches lay within the same region of the Ramachandaran plot, the matches were treated as ambiguous, and no prediction of the phi and psi values of that residue was made.

2.13.3 Structure Calculation using CNS

The program Crystallography and NMR System (CNS, version 1.1) was used to calculate $StpA_{91-134}$ NMR structures (Brunger *et al.*, 1998), using the NOE distance restraints and TALOS backbone dihedral predictions. NOE distance restraints from the ^{15}N-NOESY-HSQC spectrum were grouped into three categories: strong, medium and weak (corresponding to distances <2.9Å, <3.8Å and <5.5Å, respectively), based on NOE cross-peak intensities. Restraints from the ^{13}C-NOESY-HSQC spectrum were grouped into two categories: strong and weak (corresponding to distances <2.5Å and <5.5Å, respectively). Methyl, aromatic and methylene groups with unresolved stereospecificity were replaced with the appropriate pseudoatoms. $StpA_{91-134}$ structure calculations were performed using the simulated annealing protocols included in CNS using 'the protein-allhdg-ucl' parameters. A high temperature annealing stage was carried out on an extended structure of $StpA_{91-134}$, followed by a slow-cool annealing stage, with both stages using torsion angle dynamics. A final slow-cool annealing stage was carried out using Cartesian molecular dynamics. 20 models of $StpA_{91-134}$ with the lowest NOE energies were retained as final structures.

2.13.4 Automated NOE Assignments with ARIA

NOE resonance assignment during the process of protein structure determination by NMR is made difficult by two factors: one is the assignment of ambiguous NOEs, and the other overlapping of different NOE signals due to limited spectral dispersion. To-

gether these two problems generally reduce the number of useful NOEs and degrade the quality of the restraints, leading to lower 'resolution' structures. For example, failure to resolve the identity of an ambiguous NOE may result in discarding that NOE from further analysis. If the multiplet structure of an overlapping NOE is not recognised, for example, if a doublet NOE is mistakenly identified as a single NOE, this can lead to reduction in the total number of true restraints and incorrect estimations of the magnitude of restraints. These factors can significantly increase the difficulty of assigning NOE cross-peaks manually.

Ambiguous Restraints for Iterative Assignment (ARIA v.1.2) is a program that automates the process of NOE assignment and structure calculation (Linge *et al.*, 2001; Nilges and O'Donoghue, 1998; Nilges *et al.*, 1997). Many aspects of structure determination by NMR are automated in ARIA, such as NOE peak intensity calibration, assignment of stereospecific groups, elucidating ambiguous assignments, and deconvolving overlapping multiplet NOE peaks into their constituent NOEs. The structure calculation is performed automatically using the CNS program. Before each iteration of ARIA, NOE peak intensities are recalibrated resulting in an increasing accuracy of distance restraints. Unassigned chiral groups are treated as 'floating assignments', where ARIA swaps the chemical shift assignments of the members of the chiral group during structure calculation until a particular stereospecific assignment that best fits the NOE data is found.

The ^{13}C-NOESY-HSQC, ^{15}N-NOESY-HSQC cross-peak intensities and TALOS backbone dihedral predictions were used as inputs to ARIA. Default settings were specified in the run.cns parameter file (revision 2.10), with water refinement used in the final iteration: exceptions to the default are detailed in Table 2.5.

	^{13}C-NOESY-HSQC	^{15}N-NOESY-HSQC
Tolerance, heteronuclear dimension (ppm)	0.2	0.2
Tolerance, direct ^{1}H dimension (ppm)	0.01	0.01
Tolerance, indirect ^{1}H dimension (ppm)	0.05	0.04
Magnetic field strength (MHz)	800	500
Mixing time (ms)	80	150

Iteration	Ambiguous cut-off	Violation tolerance
0	1.01	1000.0
1	0.9999	1000.0
2	0.999	2.0
3	0.999	1.0
4	0.99	0.5
5	0.96	0.2
6	0.94	0.1
7	0.92	0.1
8	0.9	0.1

Table 2.5
Settings used for the structure calculation of $StpA_{91-134}$ using ARIA.

2.13.5. Structure Viewers and Analysis

Procheck is a program that can compare an ensemble of structures with the set of restraints used to calculate them, giving detailed information about all aspects of the structures (Laskowski *et al.*, 1996). For example, the backbone dihedral angles of all the residues in the ensemble can be plotted on a Ramachandran plot, or a particular distance restraint can be compared with the actual distance between a pair of ^{1}H nuclei in the ensemble of structures. There are many features in Procheck, of which a few will be discussed here.

AQUA is a set of programs used for the analysis of the quality of biomolecular structures solved by NMR (Doreleijers *et al.*, 1999; Laskowski *et al.*, 1996). The AquaCompl module of AQUA was used to determine the NOE restraint list completeness generated using either the manual or automatic method of $StpA_{91-134}$ structure determination. AquaCompl was accessed through the World-Wide Web (http://urchin.bmrb.wisc.edu/~jurgen/aqua/), using default settings, i.e. intra-residue contacts included, standards sets of observable atoms used, 4 Å maximum distances for experimental and model-based shells.

PyMOL was used to view protein structure files and used to generate illustrations (Delano, 2004). VMD was used to view protein structures and calculate RMSD values based on the algorithm described by Kabsch (Humphrey *et al.*, 1996; Kabsch, 1978).

2.13.6 Investigation of the interaction between $StpA_{91-134}$ and DNA by NMR

The binding of [^{1}H, ^{15}N]-labelled $StpA_{91-134}$ to a 20-mer DNA duplex (Oswel) was investigated by NMR. The concentrations of two DNA strands, of sequence CTGCACTTTAAAAAGACGTC and GACGTCTTTTTAAAGTCCAG, were estimated using ε_{260} values of 201.1 $mM^{-1}cm^{-1}$ and 197.1 $mM^{-1}cm^{-1}$, respectively. An equimolar mixture was heated to 100°C and slowly cooled to room temperature. A UV temperature scan of the DNA duplex revealed a single cooperative transition of T_m of 59 °C. Samples containing 100 μM [^{1}H, ^{15}N]-labelled $StpA_{91-134}$, 10 mM sodium phosphate pH 7.0 and 100 mM NaCl were prepared, with the concentration of DNA varied between 10 μM and 100 μM. [^{1}H, ^{15}N]-HSQC spectra of each sample were acquired at 25°C on a 500 MHz Varian NMR spectrometer.

2.14 Isothermal titration calorimetry (ITC)

ITC measures the heat associated with a binding event, which is almost universal to all binding reactions, thus can be applied to the majority of applications, without the need to derivatise the reactants beforehand. As ITC measures heat directly, it allows determination of all binding parameters such as the binding constant *K*, the heat of binding ΔH, entropy ΔS, and stoichiometry *n*. These factors, combined with the short experimental equilibration times, make ITC a powerful technique for the measurement of binding parameters (Wiseman *et al.*, 1989).

The VP-ITC (Microcal) consists of a reference cell and a sample cell housed within an adiabatic shield. A small constant current is supplied to heat the reference cell, whilst a feedback mechanism connected to a heater on the sample cell strives to keep the difference in temperature between the sample and reference cell (ΔT_1) to zero. Similarly, a feedback mechanism maintains a constant temperature of the adiabatic shield. Sample in then injected in increments into the sample cell by a syringe, which has an impeller attached to the end to ensure rapid mixing of any injected sample. When an exothermic reaction occurs, less current is required to nullify ΔT_1, whereas

with an endothermic reaction, more current is supplied to nullify ΔT_1. The heats of reaction are readily obtained by the summation of the current required to nullify ΔT_1.

Calf thymus DNA (ctDNA, from Sigma) was prepared in order to analyse the interaction of H-NS and StpA with DNA by ITC. 1 g of ctDNA was suspended in 100 ml buffer (20 mM Tris-HCl pH 7.0, 0.5 mM EDTA, 250 mM NaCl) and sonicated exhaustively (approximately 100 cycles of 1 minute sonication pulses followed by a rest period of 1 minute) using a Soniprep 150 (Sanyo MSE). When analysed by 1.5 % agarose gel electrophoresis, comparison of the sonication products with the 100 bp DNA ladder (NEB) suggested the majority of the ctDNA was between 100 bp to 500 bp in length (data not shown).

Samples of $StpA_{FL}$, $H\text{-}NS_{FL}$ and ctDNA were placed in separate dialysis tubing and dialysed against 20 mM sodium phosphate buffer pH 7.0, 300 mM NaCl, 0.1 mM EDTA and were degassed. ITC measurements were carried out using a VP-ITC (Microcal). Protein samples were placed in the ITC cell, with ctDNA placed in the injection needle. All measurements were carried out at 25°C unless stated otherwise. 18 injections were carried out, with a sample stirring speed of 300 rpm. The resulting data was analysed Excel (Microsoft) and fitted to models using Origin v5.0 (Microcal).

2.14.1 'Independent binding sites' model

ITC data was fitted according to the 'independent binding sites' model (Wiseman *et al.*, 1989):

$$\frac{1}{V_0}\left(\frac{dQ}{dX_t}\right)=\frac{\Delta H}{2}\left(1-\frac{\left(\frac{X_t}{nM_t}\right)+\frac{1}{nKM_t}-1}{\sqrt{\left(\frac{X_t}{nM_t}\right)^2-2\left(\frac{X_t}{nM_t}\right)\left(1-\frac{1}{nKM_t}\right)+\left(\frac{1}{nKM_t}+1\right)^2}}\right)+C \qquad \text{Equation 2.5}$$

M and X represent the protein ($H\text{-}NS_{FL}$) and ligand (DNA, bp) concentrations, respectively. X_t is the total concentration of X (M) and M_t is the total concentration of M (M). The stoichiometry of binding is n, (moles of X per mole of M). ΔH is the molar enthalpy of binding of per mole of X (cal M^{-1}). V_0 is the ITC cell volume (L). The

heat change dQ is measured directly in the ITC experiment (μcal). C is the molar heat of dilution of X (cal M^{-1}).

2.14.2 'Lattice-ligand' model

Detailed in this section is the method of converting ITC data into 'v' and 'L', the binding density and the concentration of free ligand, respectively, to allow fitting of the data to the 'lattice-ligand' model described by McGhee and von Hippel, 1974. M and X represent the protein (H-NS_{FL}) and ligand (DNA, bp) concentrations, respectively.

The heat of dilution of X per injection in the ITC cell, dQ_{dil}, is as follows:

$$dQ_{dil} = C \cdot V_0 \cdot dX_t \qquad \text{Equation 2.6}$$

where C is the molar heat of dilution of X (cal M^{-1}), V_0 is the ITC cell volume (L) and dX_t is the change in concentration of X per injection (M). dQ_{dil} was subtracted from the heat measured per injection, dQ, to remove the heat of dilution of X from the data.

The total heat content of within the ITC cell, Q, was calculated according to the following relationship (from the Microcal ITC data analysis handbook version 7):

$$Q_i = \frac{\left(dQ_i + Q_{i-1} - \frac{V_i \cdot Q_{i-1}}{2V_0} \right)}{1 + \frac{V_i}{2V_0}} \qquad \text{Equation 2.7}$$

where Q_i is the total heat associated with all injections up to and including injection i (cal), dQ_i is the incremental heat measured directly by the ITC apparatus at injection i (cal), Q_{i-1} is the total heat associated with injection i-1 (cal), V_i is the injection volume (L), and V_0 is the ITC cell volume (L).

The total heat content of within the ITC cell, Q, was then related to the fractional occupancy of M, θ, as follows:

$$Q = n \cdot \theta \cdot M_t \cdot \Delta H \cdot V_0 \qquad \text{Equation 2.8}$$

where n is the stoichiometry of binding (moles of X per mole of M), M_t is the total concentration of M (M) and ΔH is the molar heat of binding of X (cal M^{-1}).

The concentrations of free and bound macromolecule M was then determined:

$$M_{free} = \theta \cdot M_t \qquad \text{Equations 2.9}$$

$$M_{bound} = (1 - \theta) M_t$$

where M_{free} and M_{bound} are the concentrations of free and bound macromolecule M, respectively (M).

The binding density 'ν' (in moles of bound ligand per mole of lattice residue), and concentration of free ligand 'L' (M) was calculated as follows:

$$\nu = \frac{M_{bound}}{X_t} \qquad \text{Equations 2.10}$$

$$L = M_{free}$$

The plot of ν/L with respect to ν was then fitted to the following equation which describes cooperative binding of large ligands to a homogenous lattice (McGhee and von Hippel, 1974):

$$\frac{\nu}{L} = K(1 - n\nu)\left(\frac{(2\omega + 1)(1 - n\nu) + \nu - R}{2(\omega - 1)(1 - n\nu)}\right)^{n-1}\left(\frac{1 - (n+1)\nu + R}{2(1 - n\nu)}\right)^2 \qquad \text{Equation 2.11}$$

where n, K and ω are the stoichiometry, inherent binding constant and cooperatively parameter, respectively.

2.15 His-tag pull down assay

1 ml samples containing 100 μM His-tagged protein and 100 μM untagged protein were prepared, buffered in 20 mM sodium phosphate pH 7.0, 500 mM NaCl and 0.1 mM EDTA. Each of these samples was stored at either 4°C or at 37°C, for either one hour, for 15 hours or for 40 hours, giving a total of six different storage conditions. One identical control sample was prepared and subjected to the following protocol: the sample was dialysed into 6 M guanidine hydrochloride buffered with 20 mM sodium phosphate pH 7.0, heated at 55°C for one hour, and then serially dialysed three times into 20 mM sodium phosphate buffer pH 7.0, 500 mM NaCl, 0.1 mM EDTA.

All samples were then loaded onto a TALON spin column (BD sciences) pre-equilibrated with 20 mM sodium phosphate pH 7.0, 500 mM NaCl and 0.1 mM EDTA. The columns were gently mixed for 5 minutes. The spin columns were then centrifuged at 700 g for 2 minutes and the flow-though collected. The column-wash step was carried out by the addition of 1 ml of wash buffer (20 mM sodium phosphate pH 7.0 and 500 mM NaCl) and gently mixed for another 5 minutes. The wash buffer was removed by centrifugation at 700 g for 2 minutes. The wash step was repeated 5 times. For elution, 1 ml of elution buffer (20 sodium phosphate pH 7.0, 500 mM NaCl and 100 mM EDTA) was added to the column and the column gently mixed for 5 minutes. The eluted sample was collected by centrifugation of the column at 700 g for 2 minutes, and analysed by 12 % SDS-PAGE.

Chapter 3

The determination of the solution structure of $StpA_{91-134}$ using NMR

3.1 Introduction

In this chapter the studies of $StpA_{91-134}$ using NMR spectroscopy are presented. A brief introduction to aspects of NMR spectroscopy relevant to the field of structural biology is given. Methods of spectral assignment and structure determination by NMR are illustrated using $StpA_{91-134}$ as an example. The complete assignment and data from $StpA_{91-134}$ are presented, and results discussed. Finally an investigation of the interaction between $StpA_{91-134}$ and DNA by NMR is described.

3.1.1 Theory of NMR

Some basic principles of NMR are presented in this section, discussed in context to application of NMR to biological macromolecules. The NMR phenomenon arises from a quantum mechanical property inherent to fundamental particles called spin. An atomic nucleus has a total spin arising from the constituent protons and neutrons. The closest analogy to nuclear spin is the geometric spin or the angular momentum of a rotating object in classical mechanics. However the nucleus itself does not physically rotate as a result of nuclear spin: rather, it is an intrinsic property of the nucleus. Unlike other properties such as mass and charge, nuclear spin has virtually no effect on the macroscopic chemical and physical behaviour of substances.

Nuclear spin, I, can have integer or half-integer values, for example 0, 1/2, 1, 3/2, 2, 5/2 and so on. The nuclei that are usually of interest to biological NMR, ^{1}H, ^{13}C and ^{15}N, all have nuclear spin of ½, to which can be applied the analogy of a freely rotating bar magnet with a magnetic moment of μ. When placed in the static magnetic field B_0 of a NMR spectrometer, the axis of rotation of the magnetic moment aligns itself parallel (α-state, with a positive μ) or anti-parallel (β-state, with negative μ) with B_0. The magnetic moment then precesses around B_0, as shown in Figure 3.1 (a). These two states are known as Zeeman states. Note that the angle between the nuclear magnetic moment at any particular point in time and the axis of precession is a quantised property and remains constant. The population of nuclei occupy the two Zeeman states, with nuclei preferentially occupying the lower energy α-state at thermal equilibrium. At experimentally achievable field strengths, the difference between the energy states α and β is very small, leading to a very small bias in spin polarisation in favour of the α-state at thermal equilibrium.

At equilibrium the transverse component (the x, y component) of the magnetic moment is essentially zero, as the phases of all the different precessing nuclei are random, as shown schematically in Figure 3.1 (b). Summation of all the magnetic moments of the nuclei result in a bulk magnetisation M_Z aligning with the field of strength B_0 with no net transverse component (i.e. there is no net magnetisation in the x, y planes), as shown in Figure 3.1 (c).

Once aligned to the static magnetic field, the nuclear spin precesses around the z-axis at a rate called the Larmor frequency, ω_0. NMR experiments are performed by the application of B_1 - an oscillating magnetic field, i.e. a radiofrequency (RF) pulse - to the nuclei, perpendicular to B_0, to manipulate the bulk magnetisation of the nuclear spins. The B_1 pulse produces torque on the bulk magnetisation around the B_1 pulse vector, causing the bulk magnetisation to rotate around that vector. The degree of rotation is determined by the strength and the duration of the B_1 pulse. 90° pulses (or $\pi/2$) and 180° pulses (or π) are commonly used in NMR experiments; Figure 3.2 shows the effects of a $\pi_x/2$ pulse and a π_x pulse on the bulk magnetisation vector M_Z.

Any rotating magnetic moment, such as the nuclear spins precessing in the static field B_0, generates a rotating magnetic and electric field, which in turn generates an electric

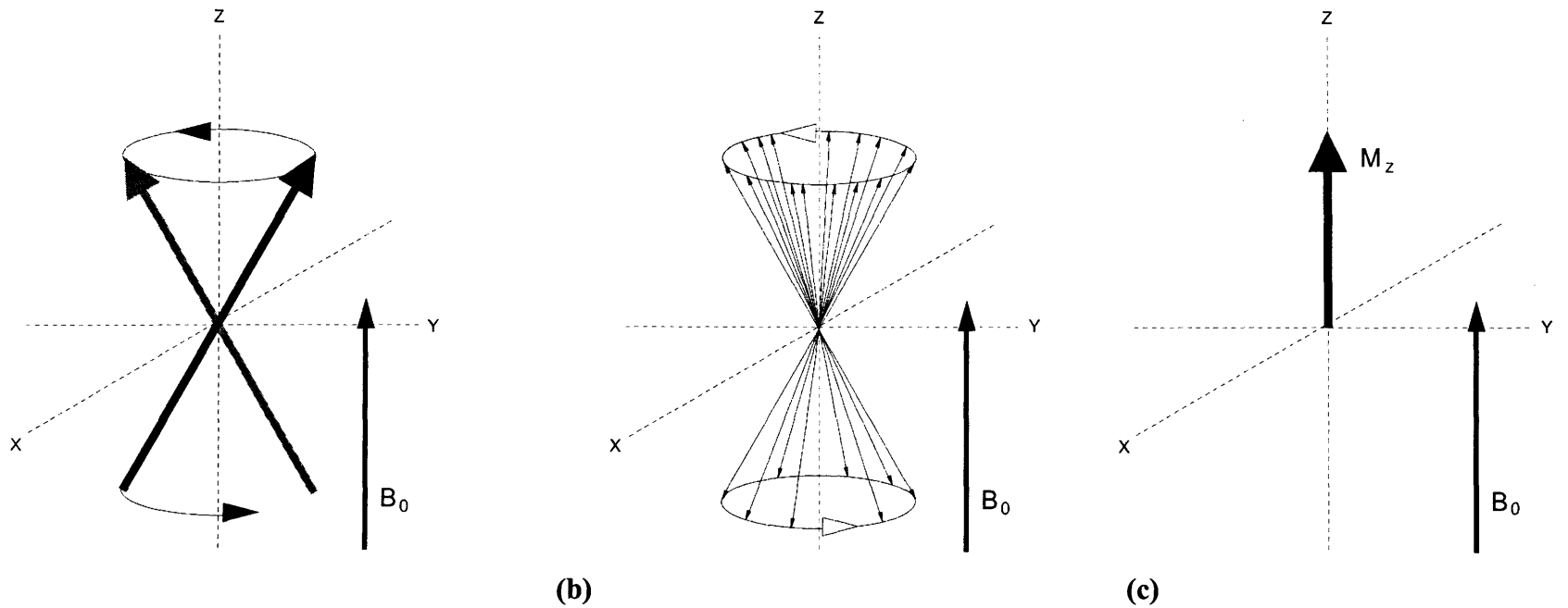


Figure 3.1

(a) A figure showing a vector model of the nuclear magnetic moment of a nucleus with spin ½. In the presence of the static field B_0, the nuclear magnetic moment aligns parallel or anti-parallel (in this case, parallel) and precesses in a plane perpendicular to B_0.

(b) A figure showing a vector model of several superimposed nuclear magnetic moments of nuclei with spin ½. At thermal equilibrium, more nuclear magnetic moments align with the static field B_0 than against the field. (The proportion of the nuclear magnetic moments aligning with and against is shown schematically.) The phases of all of the individual nuclear magnetic moments are random.

(c) A figure showing the vector model of the bulk magnetisation M_Z, aligning with the static field B_0. The bulk magnetisation represents the summation of all the individual vector representations of the nuclear magnetic moments shown in (b)

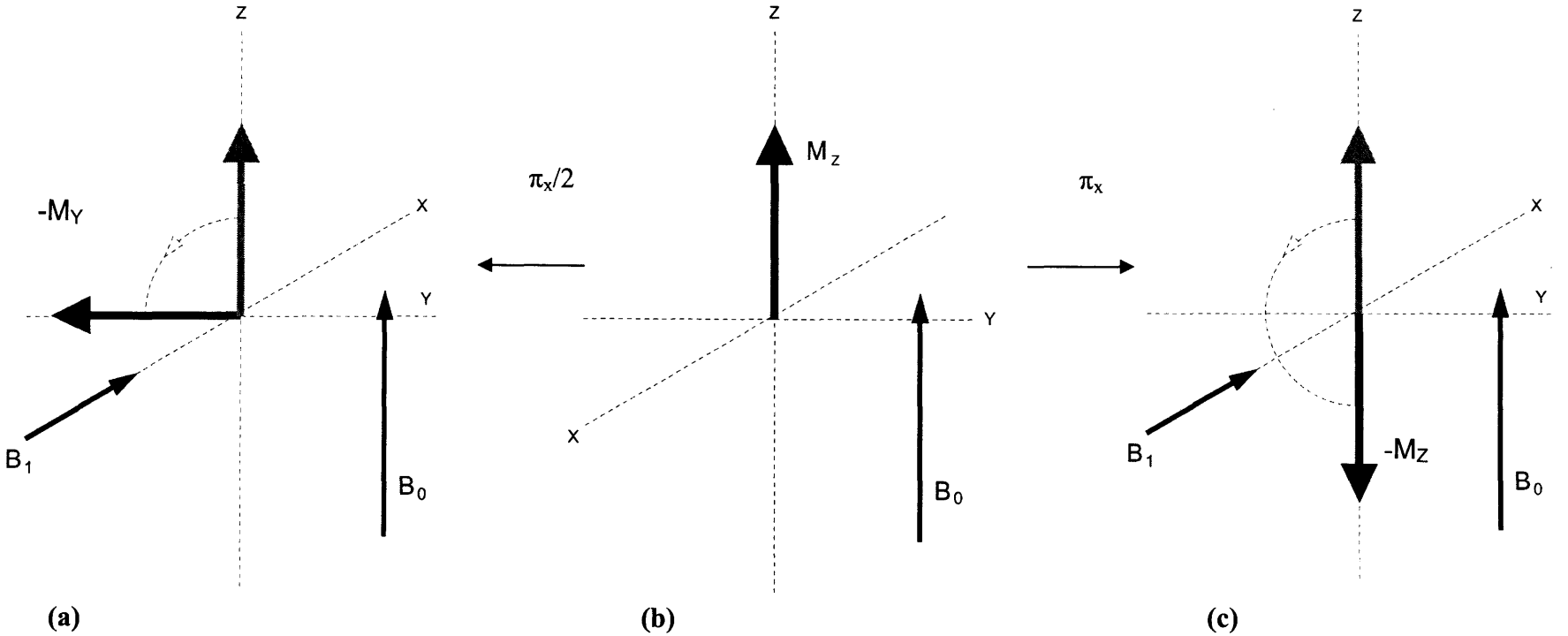


Figure 3.2

A figure showing the effect of a $\pi_x/2$ pulse and a π_x pulse on the bulk magnetisation vector M_Z. At thermal equilibrium the nuclear magnetic moments of spin ½ preferentially align with the static field B_0, resulting in the bulk magnetisation vector M_Z, as shown in Figure 3.2 (b). When a $\pi_x/2$ RF pulse is applied along the *x*-axis, the bulk magnetisation rotates 90° onto the transverse plane, giving rise to coherence, shown as the bulk magnetisation vector $-M_Y$ in Figure 3.2 (a). If, instead, a π_x pulse is applied along the x-axis, the bulk magnetisation vector M_Z rotates 180°, aligning against B_0, shown in Figure 3.2 (c).

current in any nearby coils. When there is no transverse component in the bulk magnetisation, for example M_z, as shown in Figure 3.2 (a) or $-M_z$, as shown in Figure 3.2 (c), no signal is detected. When there is a transverse component to the bulk magnetisation, for example after a $\pi_x/2$ pulse, as shown in Figure 3.2 (b): under the influence of B_0, this bulk magnetisation precesses around the *z*-axis at the Larmor frequency ω_0: the transverse component of the rotating magnetic and electric field of the precessing bulk magnetisation generates a signal which is recorded. (In reality, in NMR signals are recorded in the rotating frame, where the signals are viewed in a reference frame that rotates around the *z*-axis with a frequency of ω_0. Thus any bulk magnetisation rotating around the *z*-axis with frequency ω_0 would look static in the rotating frame).

Thus the NMR technique can be summarised as follows: first, nuclear spins align in a large static magnetic field B_0, precessing around the B_0 field vector at the Larmor frequency. This is followed by the NMR experiment: the nuclear spins are rotated by the use of a radiofrequency pulse B_1 around the B_1 pulse vector. After the application of the radiofrequency pulse, the nuclear spins precess around the B_0 field, generating a rotating magnetic and electric field that is amplified and recorded.

The signal from a NMR spectrometer is known as the free induction decay (FID), which is an oscillating function that decays exponentially, acquired as a function of time. The exponential decay results from relaxation processes as the nuclear spins return to thermal equilibrium (relaxation is discussed further in Section 3.1.3). A simple representation of an FID is shown in Figure 3.3. The time domain of the NMR signal is converted into the frequency domain by Fourier transformation, giving rise to a one dimensional (1D) NMR spectrum. The Fourier transform of the FID oscillating at the Larmor frequency ω_0 gives an absorption Lorentzian function with the centre of the peak at ω_0.

Whilst the applied static field B_0 can be made to be very homogenous, small differences in local environment mean that nuclei experience different magnetic environments. The nuclei themselves induce a magnetic field; hence the local magnetic environment of a particular nucleus depends on the chemical environment and surrounding nuclei as well as B_0. For example, ^1H nuclei in a protein situated in different parts of the structure can experience very different local chemical environments, as a result

of their proximity to aromatic side chains, chelated metal ions, exposure to water molecules and so on. The slight shift in the Larmor frequency of any particular nucleus as a result of a subtle change of the magnetic field exerted by the local environment is known as the chemical shift. As the Larmor frequencies are dependent on the magnetic field strength of the NMR spectrometer, chemical shift is usually expressed as ratios of the chemical shift frequency to the magnetic field strength (in units of parts per million) to allow comparisons between different NMR spectrometers. The NMR spectrum can be processed by treatment of the FID with mathematical algorithms (see Section 2.13.1), resulting in improved resolution and quality.

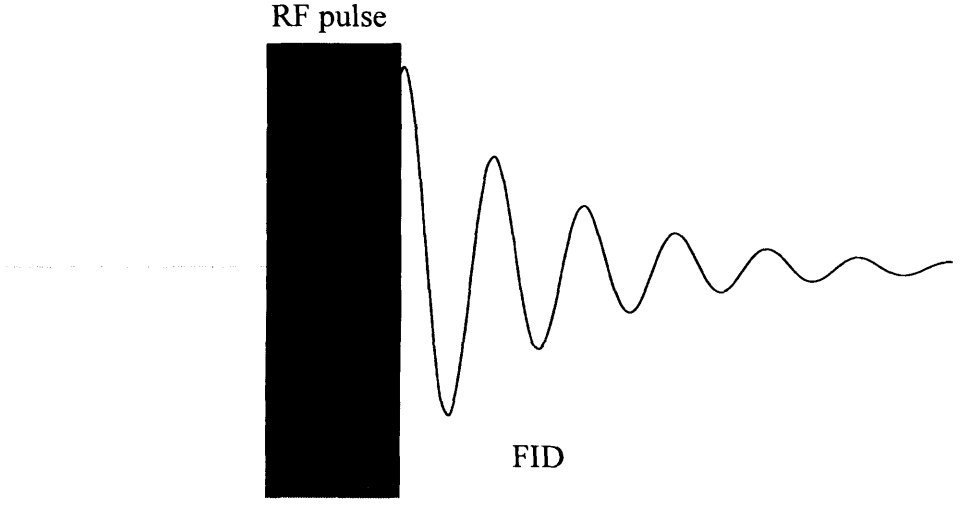


Figure 3.3
A diagram showing a basic NMR experiment: a RF pulse, represented as the grey block in the figure above, causes coherence in the nuclear spins. As soon as the RF pulse finishes, the nuclear spins precess in the transverse plane, inducing an alternating current in the detection coils which is recorded as the FID. The rate of precession of the nuclear spins determines the oscillating component of the FID, whilst relaxation (discussed in Section 3.1.3) determines the exponentially decaying component of the FID.

3.1.2 Multidimensional NMR

There are many limitations to the NMR technique when applied to the study of proteins; even with small proteins, there are hundreds of ^{1}H nuclei resonances, the majority of which must be assigned before any structural analysis can be initiated. In the limited spectral width of the ^{1}H nuclei resonances, this can lead to severe signal overlap and degeneracy. When analysing larger proteins, not only does the number of ^{1}H nuclei resonances increase accordingly, but the increasing molecular mass of the protein leads to unfavourable relaxation conditions that lead to spectral line-width broadening (see section 3.1.3). These limitations dictate that 1D NMR can only be realistically applied to the study of peptides.

The development of two-dimensional (2D) and three-dimensional (3D) NMR techniques alleviated some problems associated with studying larger proteins, and extended the limit of molecular mass using standard NMR techniques to about 30 kDa. In homonuclear 2D ^{1}H-based experiments, the NMR spectrum is dispersed into the second dimension, and can resolve problems associated with spectral overlap (Bax, 1989; Wuthrich, 1986). Higher dimensional techniques were developed by the introduction of enrichment in ^{13}Carbon (^{13}C) and ^{15}Nitrogen (^{15}N) isotopes in proteins by taking advantage of heteronuclear couplings, allowing spectra to be further dispersed into a third dimension. Additional information can be obtained from the heteronuclear spectra, such as the determination of secondary structure using ^{13}C chemical shifts values.

3.1.3 NMR relaxation

Relaxation is the process by which thermal equilibrium is established in a population of nuclei in the magnetic field. The relaxation behaviour of a protein gives useful insights into its molecular dynamics by providing information about global and local motions of the protein structure, as well as providing an estimate of the shape and molecular mass of the protein.

There are two mechanisms through which this can happen: longitudinal relaxation (T_1) and transverse relaxation (T_2). In the presence of a static field B_0, the magnetic moment of a nuclear spin has a tendency to align with B_0, which represents a lower energy state (the α-energy state), than against it (the β-energy state). Longitudinal

relaxation refers to the process of the nuclear spin bulk magnetisation aligning with the static field B_0 to establish a thermal equilibrium (or to the loss of the bulk magnetisation when B_0 is turned off).

Transverse relaxation refers to the process by which phase coherence of the nuclear spins is lost and precessing spins are spread uniformly across the transverse plane. The following example illustrates transverse relaxation: when the nuclear spins align in the static field B_0, a net magnetic moment is created aligned with the z-axis, giving a bulk magnetisation M_z, as shown in Figure 3.2 (b). As there is no phase coherence of the spins, there is no net transverse component. The bulk magnetisation M_Z is then disturbed from equilibrium by a $\pi_x/2$ pulse along B_1, giving a bulk magnetisation $-M_Y$ as shown in Figure 3.2 (a). The bulk magnetisation is now perpendicular to B_0, and the nuclear spins are said to have transverse magnetisation. Under the influence of the static field B_0, the nuclear spins precess around the z-axis together, initially in phase. However, as explained in Section 3.1.1, the nuclei experience subtly different magnetic fields as a result of their local environment: consequently their Larmor frequencies are slightly different. Hence after a certain time period, the nuclear spins will lose their transverse phase coherence by transverse relaxation.

Transverse relaxation rates increase with increasing molecular mass of the protein. Decreasing relaxation times result in increasing line-widths of the NMR signal, such that when analysing proteins with a molecular mass greater than 30 kDa, line-widths are substantially broader. In these situations, NMR signals often overlap to the extent that the spectra are uninterpretable; consequently 30 kDa is the maximum molecular weight of proteins whose structure was solved by standard protein NMR techniques. Recently, developments in techniques in NMR such as transverse relaxation optimised spectroscopy (TROSY) and the use of liquid crystals has enabled analysis of proteins much larger than 30 kDa. However, as these techniques were not necessary for the work described in this thesis, they will not be described here.

3.2 Determination of the solution structure of $StpA_{91-134}$ by NMR

In this section the process of the determination of the NMR solution structure of $StpA_{91-134}$ is presented. This was a sequential process of sample preparation, NMR data acquisition and processing, backbone resonance assignment, side chain resonance assignment, nuclear Overhauser effect (NOE) resonance assignment, and finally structure calculation. In the context of $StpA_{91-134}$, these steps are discussed in detail. The different NMR experiments are briefly introduced, and the NMR data acquired from $StpA_{91-134}$ is used to illustrate the methods involved. Finally, analysis of the $StpA_{91-134}$ solution structure is presented.

3.2.1 Overview

A 1.5 mM sample of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ containing 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3, prepared as outlined in Sections 2.6 and 2.7, was used in all the NMR experiments detailed in this chapter.

The backbone connectivities were established and the ^{1}H, ^{15}N and ^{13}Cα resonances assigned by analysis of the triple resonance experiments HNCACB, CBCACONH and CCONH. ^{15}N-TOWNY-HSQC and HCCONH spectra were inspected to confirm backbone amide ^{1}H and ^{15}N resonances assignments.

The side chain resonance assignments were carried out by analysis of the following spectra: HNHA, HNHB, HNCACB, [^{1}H, ^{13}C]-CT-HSQC and two HCCH-TOCSYs (each with different mixing times). Aromatic side chain resonances were assigned using the [^{1}H, ^{13}C]-CT-HSQC and HCCH-TOCSY spectra with modified offsets optimised to detect aromatic side chain resonances, and a HBCD spectrum. The backbone and side chain resonance assignment list is shown in Section 3.6.

Using the chemical shift values of ^{1}Hα, ^{13}CO, ^{15}N, ^{13}Cα and ^{13}Cβ as inputs to the 'Torsion Angle Likelihood Obtained from Shift and sequence similarity' program (TALOS), φ (phi) and ψ (psi) backbone dihedral angles were estimated.

With all of the backbone and side chain resonance assignments established, NOE resonances in the ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra were assigned

manually. $StpA_{91-134}$ model structures were then calculated using the 'Crystallography and NMR System' program (CNS), with the NOE resonance assignments and the backbone dihedral angle estimates as restraints. An alternative method involving automatic NOE resonance assignment and structure calculation was also carried using the 'Ambiguous Restraints for Iterative Assignment' program (ARIA).

3.2.2 Backbone resonance assignment

A number of triple resonance experiments facilitated backbone resonance assignment and determined backbone connectivities of $StpA_{91-134}$ enriched with ^{15}N and ^{13}C. Heteronuclear experiments provided connectivities between a residue and its preceding residue by specifically correlating nuclei using a variety of one-bond and two-bond heteronuclear couplings. A selection of triple resonance experiments provided enough data to make unambiguous backbone resonance assignments relatively straightforward. As the nuclear resonances of these experiments are all detected off the backbone amide HN of residues, the [^{1}H, ^{15}N]-HSQC was used as a reference spectrum (see Section 3.2.2.1).

The backbone resonance assignments of $StpA_{91-134}$ were established using the triple resonance experiments HNCACB and CBCACONH. The HNCACB experiment correlates the $^{13}C\alpha$ and $^{13}C\beta$ nuclei of a given residue with its backbone amide HN, as well as the $^{13}C\alpha$ and $^{13}C\beta$ of the preceding residue (see Section 3.2.2.2). The CBCACONH experiment correlates only the $^{13}C\alpha$ and $^{13}C\beta$ nuclei of the preceding residue with the backbone amide HN of a given residue (see Section 3.2.2.3). Comparison of these two spectra allowed discrimination of the inter- and intra-residue $^{13}C\alpha$ and $^{13}C\beta$ resonances in the HNCACB spectrum and aided in establishing sequential connectivities between a particular residue and its preceding residue. This process was repeated from the C-terminus to the N-terminus to determine backbone connectivities of $StpA_{91-134}$. Certain residues, such as glycine, threonine, serine and alanine, have characteristic $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts, and were easily identified in the HNCACB and CBCACONH spectra. The CCONH experiment correlates the backbone amide HN of a residue with all of the ^{13}C nuclei (except the backbone carbonyl) of the preceding residue (see Section 3.2.2.4). As different amino acids have different side chain lengths, the CCONH spectrum further aided in the identification

of amino acid types. The backbone resonance assignment of $StpA_{91-134}$ was successfully completed using the HNCACB, CBCACONH and CCONH spectra.

The backbone connectivities were confirmed with the ^{15}N-TOWNY-HSQC and HCCONH experiments. The ^{15}N-TOWNY-HSQC experiment correlates a backbone amide HN with all of the intra-residue ^{1}H nuclei (see Section 3.2.2.5), and the HCCONH experiment correlates the backbone amide HN of a residue with all of the ^{1}H nuclei of the preceding residue (see Section 3.2.2.6). The backbone connectivities determined using the two sets of spectra matched, giving higher confidence to the reliability of the backbone resonance assignment.

3.2.2.1 [^{1}H, ^{15}N]-HSQC

The [^{1}H, ^{15}N]-HSQC experiment shows the backbone amide resonances of all the residues of a protein (with the exception of proline, which lacks an amide proton), as well as some of the side chain HN groups of residues such as tryptophan, glutamine, arginine and asparagine (Kay *et al.*, 1992a). Different proteins have easily distinguishable [^{1}H, ^{15}N]-HSQC spectra, sometimes referred to as a 'fingerprint', which is a sensitive qualitative indicator of the folded state of that protein. The [^{1}H, ^{15}N]-HSQC of $StpA_{91-134}$ is presented in Figure 3.4. Each cross peak that is labelled with a number and amino acid type corresponds to the backbone amide HN of that residue (unless the cross peak corresponds to a side-chain HN, in which case the cross peak is labelled appropriately).

All of the backbone amide resonances of $StpA_{91-134}$ are observed in the [^{1}H, ^{15}N]-HSQC spectrum, except Gly1, Ser2, Met3, Thr24 and Gly25. In some cases, flexible regions of structures (such as N- and C-termini or loops) of a given protein can occupy a range of different conformations. If the motional processes are fast relative to the chemical shift timescale (i.e. faster than μseconds), the local field corresponding to the flexible region of structure will average out over the different conformations, resulting in intense cross-peaks. For example, the intense cross-peak corresponding to Ile-48 is the result of such rapid motional processes. If the flexible regions of structure experience slow conformational exchange, the NMR signal is spread over the various confomers: each different conformation can yield a different set of cross peaks in the [^{1}H, ^{15}N]-HSQC, with the intensity of those cross peaks corresponding to

the proportional abundance of a given conformation. In addition, flexible regions of proteins are often exposed to solvent: the backbone amide HN proton may exchange rapidly with the solvent, thereby not allowing complete transfer of magnetisation between the backbone amide ^{1}H and ^{15}N nuclei. This rapid exchange of the backbone HN with solvent decreases the intensity of the corresponding cross peak. For instance the ^{1}H nuclei of NH_3^+ groups exchange rapidly with the solvent. Consequently the NH_3^+ groups of the N-terminus and of lysine and arginine residues of $StpA_{91-134}$ are not visible in a [^{1}H, ^{15}N]-HSQC. The residues Gly-1, Ser-2, Met-3, Thr-24 and Gly-25 of $StpA_{91-134}$ may not be visible in the [^{1}H, ^{15}N]-HSQC spectrum because of slow conformational exchange and/or rapid exchange of the backbone amide HN proton with solvent.

$StpA_{91-134}$ was expressed as a fusion protein with an N-terminal His-tag, separated in sequence by a thrombin cleavage site. The His-tag was removed using thrombin before NMR analysis; however, $StpA_{91-134}$ retained the residues Gly-1, Ser-2 and Met-3 of the tag at the N-terminal end; these three N-terminal residues are not observed in the [^{1}H, ^{15}N]-HSQC. This suggests these three residues undergo slow conformational exchange and/or their backbone amide HN protons are in rapid exchange with the solvent. As these three N-terminal residues are not found in the native StpA amino acid sequence, and are unlikely to affect further analyses of $StpA_{91-134}$, the absence of the cross peaks corresponding to these residues in the [^{1}H, ^{15}N]-HSQC spectrum is of little concern. These residues are excluded from further analysis.

Peaks corresponding to the backbone amide HN of residues Thr-24 and Gly-25 are not observed in the [^{1}H, ^{15}N]-HSQC spectrum, suggesting that these residues may reside on a loop that displays slow conformational exchange and/or are exposed to solvent. The homologous residues in the known structure of $H\text{-}NS_{90-136}$ indeed correspond to a flexible loop (Shindo *et al.*, 1995; Shindo *et al.*, 1999).

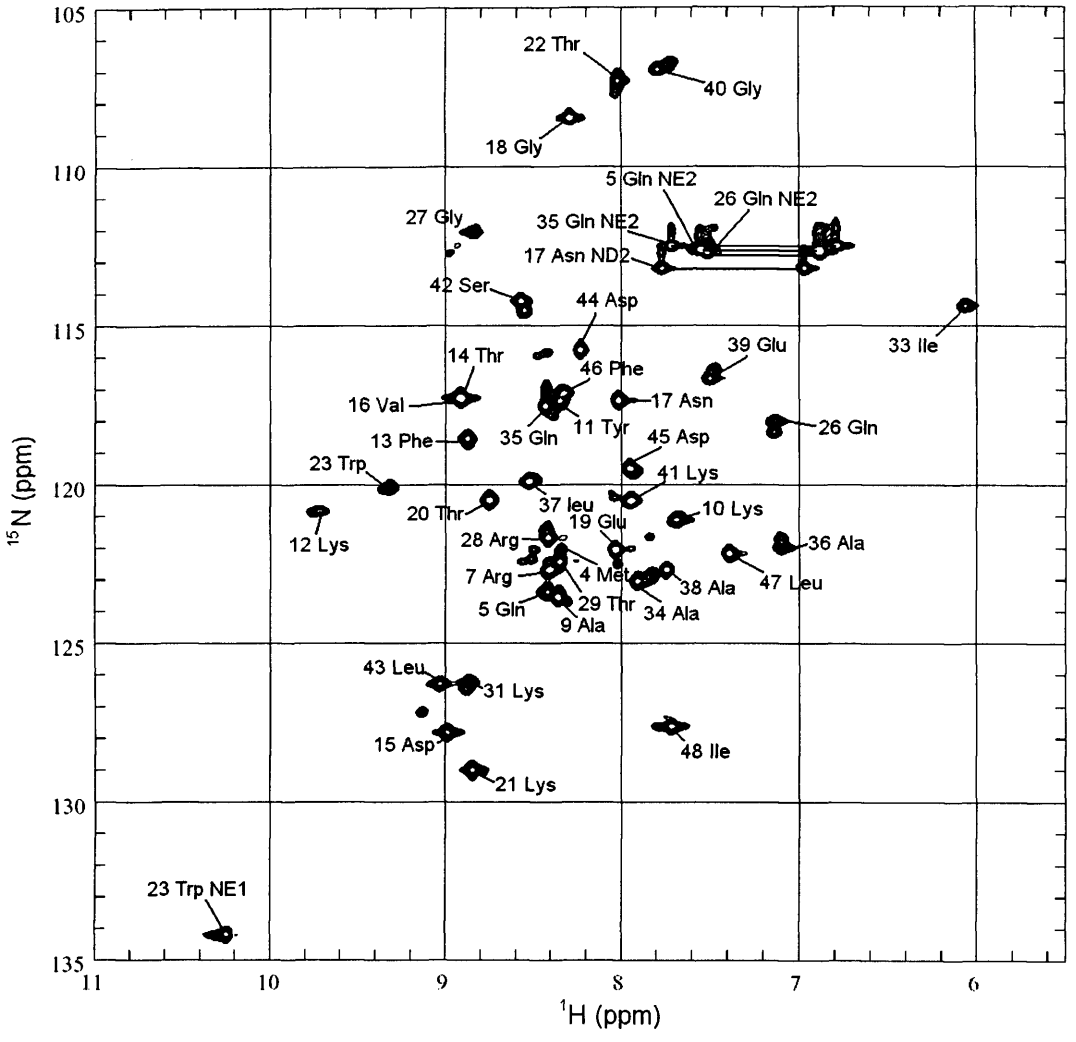


Figure 3.4
[^{1}H, ^{15}N]-HSQC spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ recorded on a 500 MHz Varian Spectrometer at 25°C. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3. Cross-peaks are labelled with the corresponding residue identity. Unlabelled cross-peaks are derived from alternate minor conformations of $StpA_{91-134}$ (as discussed in Section 3.5).

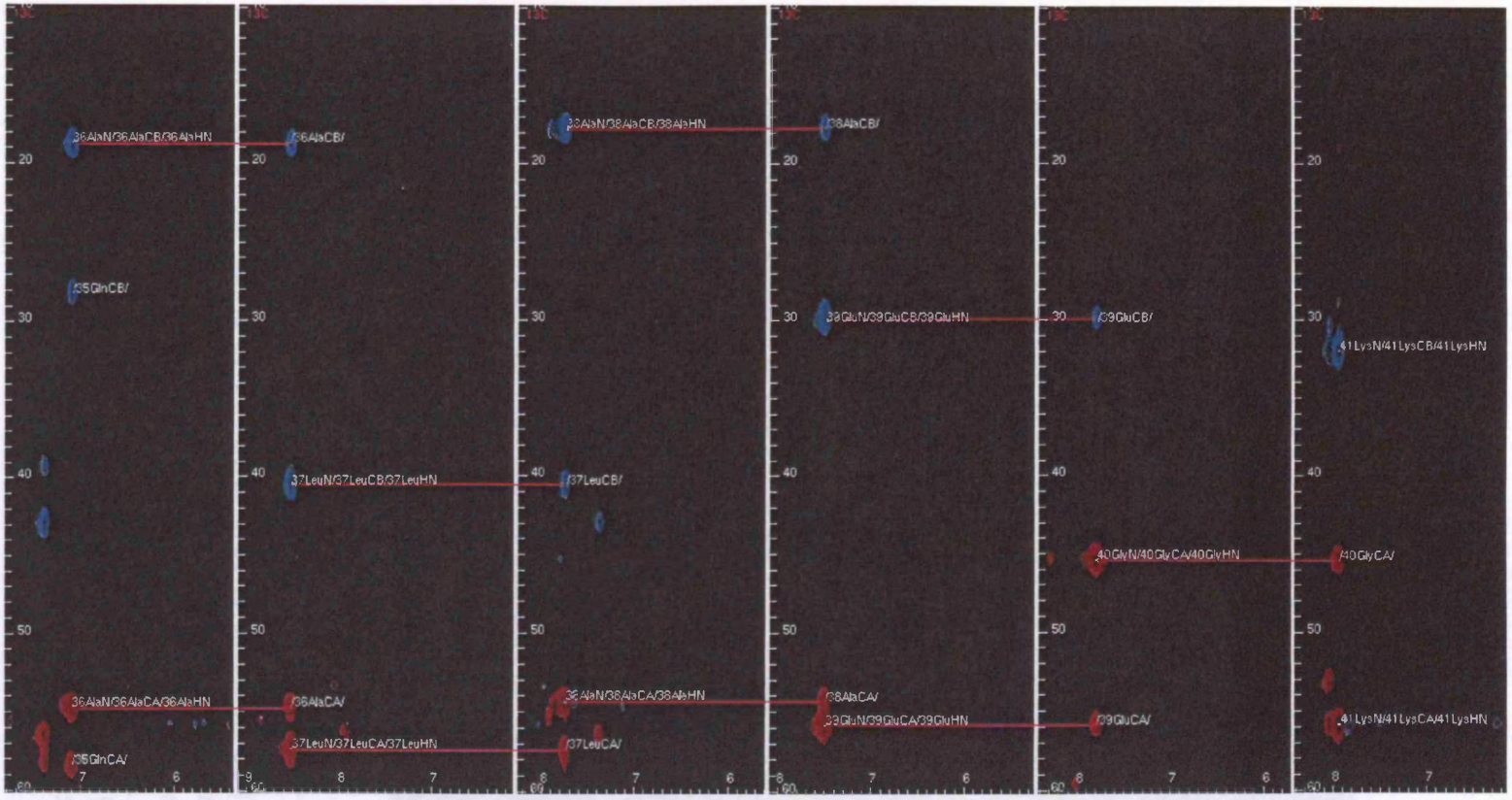


Figure 3.5

A selection of 2D slices of the HNCACB spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$, showing resonance cross peaks corresponding to residues 36 to 41 of the amino acid sequence of StpA$_{91\text{-}134}$. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3. The spectrum was recorded on a 500 MHz Varian spectrometer at 25°C.

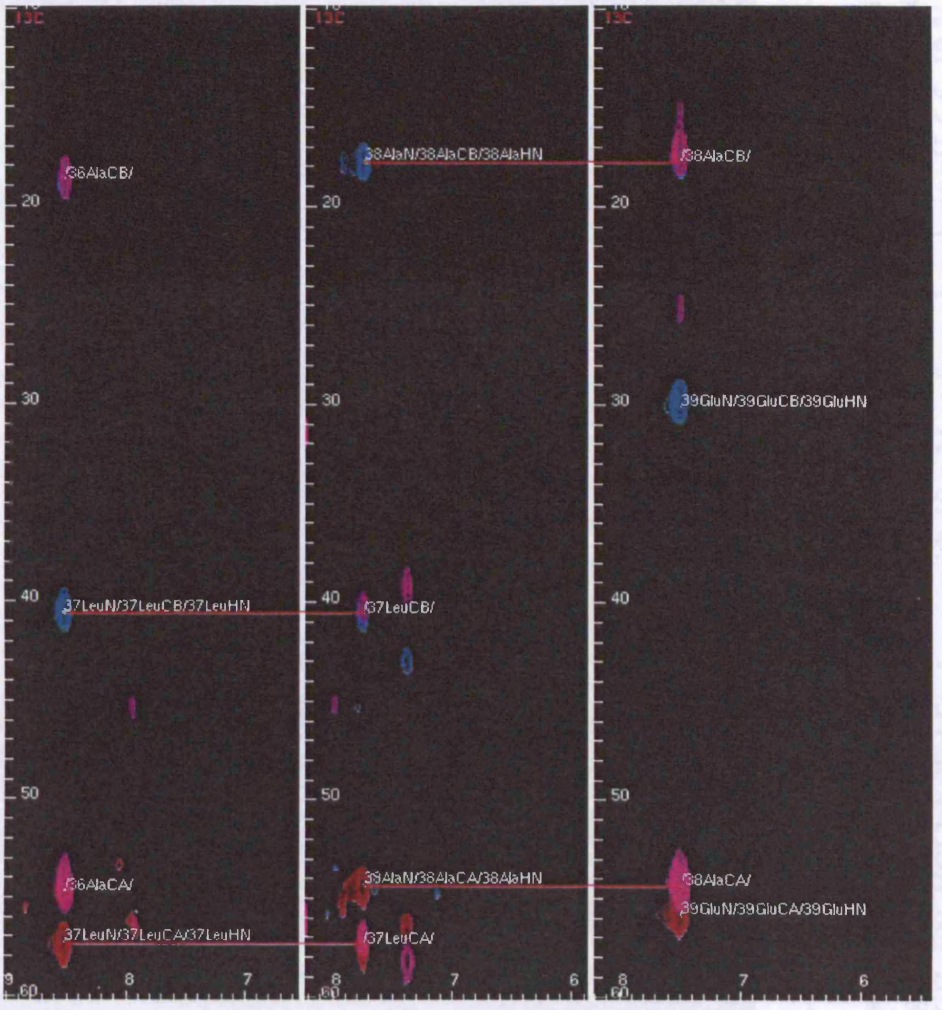


Figure 3.6
A selection of 2D slices of the CBCACONH spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ (coloured pink), superimposed on the HNCACB spectrum (coloured blue and red), showing resonance cross peaks corresponding to residues 36 to 38 of the amino acid sequence of $StpA_{91-134}$. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3. The spectrum was recorded on a 500 MHz Varian spectrometer at 25°C.

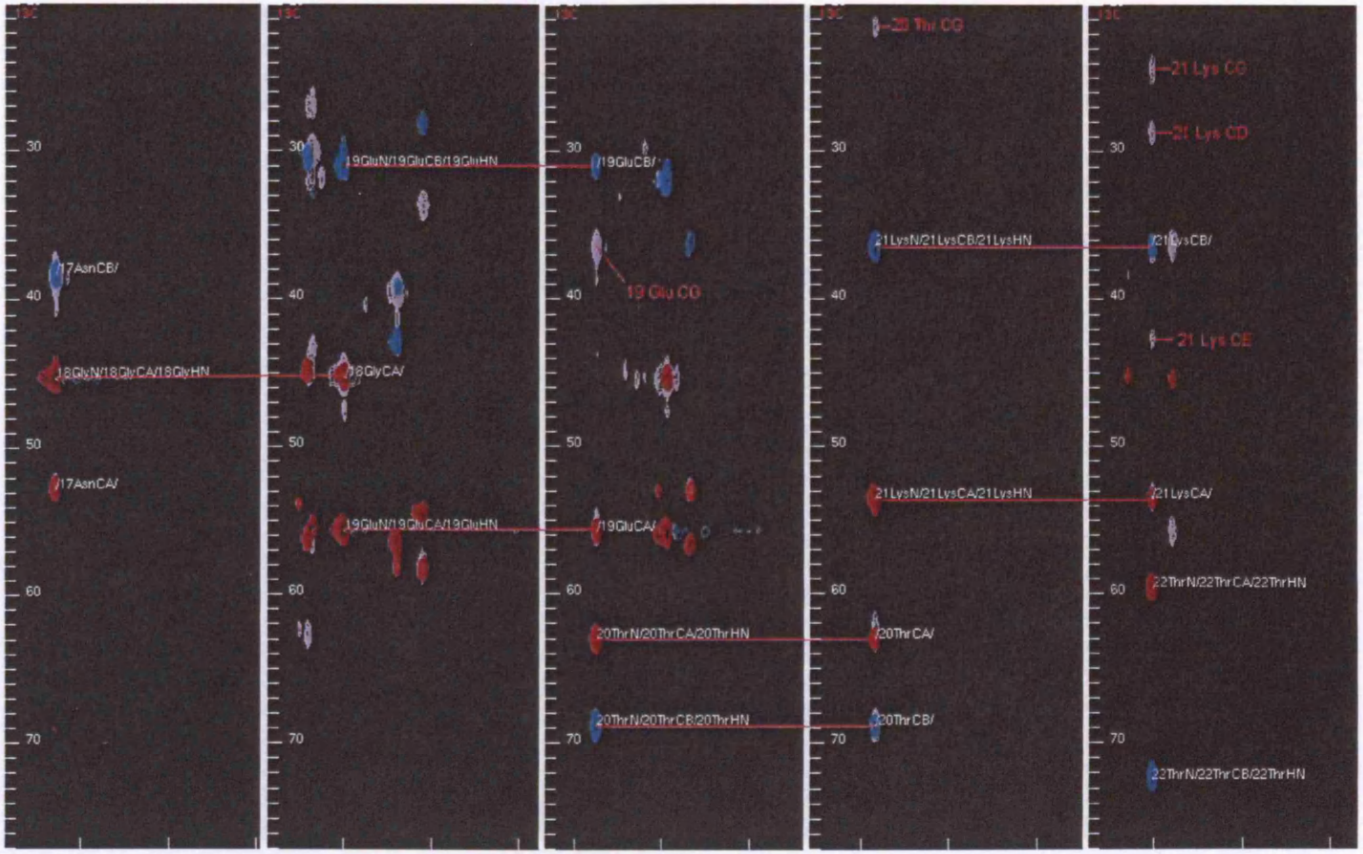


Figure 3.7

A selection of 2D slices of the CCONH spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ (coloured grey), superimposed on the HNCACB spectrum (coloured blue and red), showing resonance cross peaks corresponding to residues 18 to 21 of the amino acid sequence of StpA$_{91\text{-}134}$. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3. The spectrum was recorded on a 500 MHz Varian spectrometer at 25°C.

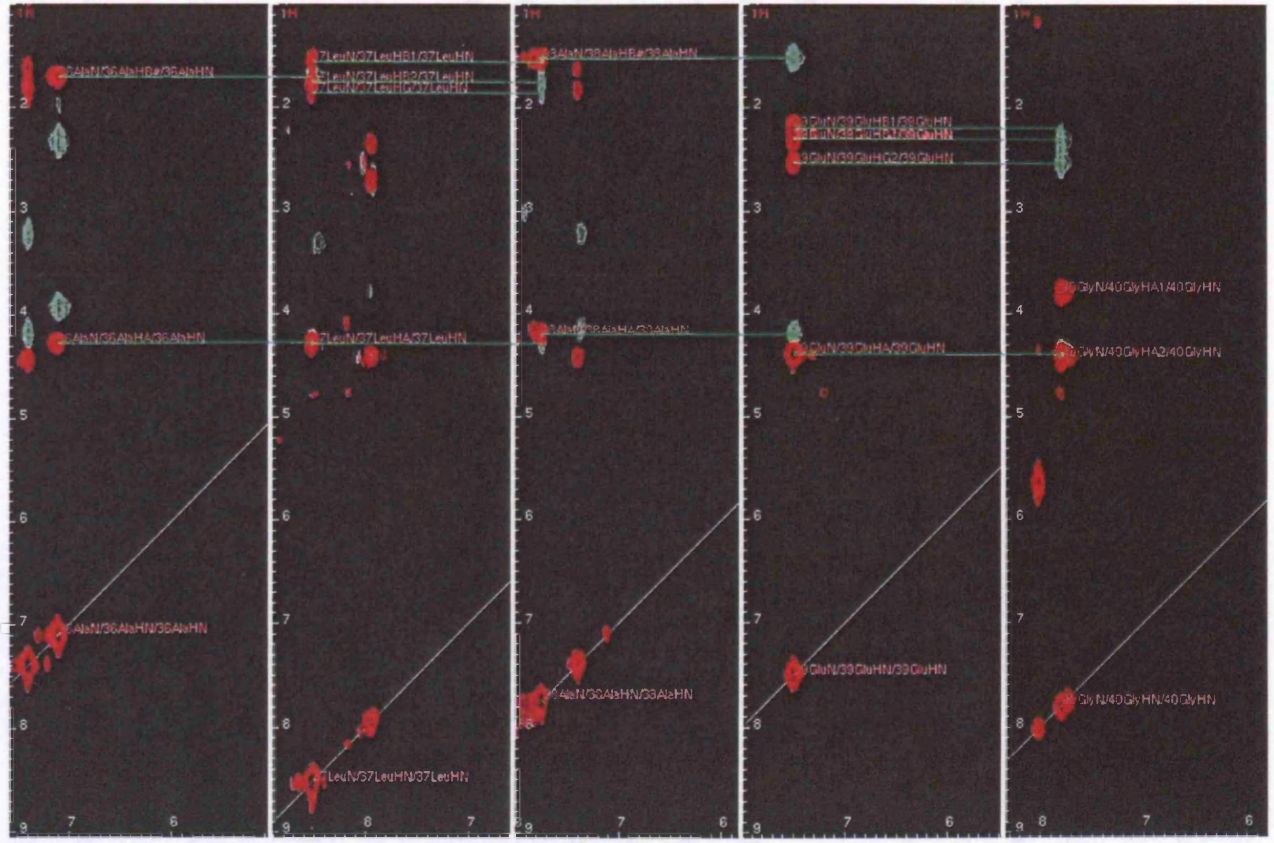


Figure 3.8

A selection of 2D slices of the [^{1}H, ^{15N}]-TOWNY-HSQC spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ (with contours coloured red) and the corresponding HCCONH spectrum (with contours coloured light green), showing resonance cross peaks corresponding to residues 36 to 40 of the amino acid sequence of $StpA_{91-134}$. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3. Both spectra were recorded on a 500 MHz Varian spectrometer at 25°C.

3.2.2.2 HNCACB

The HNCACB experiment shows correlations between the backbone amide HN and the Cα and Cβ nuclei of the same residue and those of the preceding residue (Kay *et al.*, 1992a; Yamazaki *et al.*, 1994). As well as differences in their chemical shifts, Cα and Cβ resonances of $StpA_{91-134}$ were easily distinguished by their phases, which were positive and negative, respectively. In particular, the HNCACB spectrum aided in the identification of serine and threonine residues, which have Cβ resonances at lower chemical shifts than those of Cα nuclei, in contrast to all other residues. In addition, the inter-residue resonances tended to have a lower intensity than the intra-residue Cα and Cβ resonance. Figure 3.5 shows a selection of sample 2D slices of the HNCACB spectrum, corresponding to residues Ala-36 to Lys-41. The positive Cα resonance and the negative Cβ resonance cross-peak contours are coloured red and blue, respectively. Sequential backbone connectivities are highlighted by lines.

3.2.2.3 CBCACONH

The CBCACONH experiment shows the correlation of the backbone amide HN nucleus with the Cα and Cβ resonances of the preceding residue (Kay and Muhandiram, 1994). Figure 3.6 shows a selection of 2D slices of the CBCACONH spectrum of $StpA_{91-134}$ (with cross peak contours coloured pink) superimposed on the HNCACB spectrum (as before), showing residues Ala-36, Leu-37 and Ala-38. Sequential connectivities are highlighted with lines.

3.2.2.4 CCONH

The CCONH spectrum displays correlations of the backbone amide HN with all of the aliphatic ^{13}C nuclei of the preceding residue (Grzesiek and Bax, 1993). The CCONH spectrum aided in the identification of glycine and lysine residues of $StpA_{91-134}$, the only amino acids that have one and five aliphatic carbon nuclei in their side chains, respectively. Note that in order for a carbon nucleus to be detected in this spectrum, it must have at least one ^{1}H nucleus covalently attached, therefore certain carbon nuclei resonances are not observed on the CCONH spectrum, for example, the Cγ of asparagine, the Cδ of glutamine, the Cγ of aspartate, the Cδ of glutamate, and Cζ of arginine. Figure 3.7 shows a selection of 2D slices of the CCONH spectrum (with contours coloured grey). The HNCACB spectrum (as before, with blue and red contours)

is also shown. Sets of resonances corresponding to those of Gly-18 to Thr-22 are shown, with sequential connectivities are shown as lines.

3.2.2.5 ^{15}N-TOWNY-HSQC

The 3D ^{15}N-TOWNY-HSQC experiment shows correlations between an amide HN (both the backbone amide and side chain HN) with all the intra-residue ^{1}H nuclei (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991). Samples of 2D slices of the ^{15}N-TOWNY-HSQC, with resonance contours coloured red, are shown in Figure 3.8. The backbone amide HN resonance is shown as a diagonal cross peak, which directly corresponds to a cross peak in the [^{1}H, ^{15}N]-HSQC spectrum.

3.2.2.6 HCCONH

This experiment correlates a HN group (both the backbone amide and side chain HN) with all ^{1}H nuclei of the preceding residue (Grzesiek and Bax, 1993). The identity of the residue preceding any given residue can straightforwardly be determined by comparing the HCCONH spectrum with the ^{15}N-TOWNY-HSQC spectrum. Repeating this process iteratively allows backbone connectivities to be determined from the C-terminus to the N-terminus. Samples of 2D slices of the HCCONH spectrum, with resonance contours coloured light green, are shown in Figure 3.8. Viewing the ^{15}N-TOWNY-HSQC and the HCCONH spectra together provides an alternative method to using the triple resonance experiments for assigning the backbone resonances. The sequential connectivities between the residues are highlighted.

3.2.3 Side chain resonance assignments

Once the backbone resonance assignments of $StpA_{91-134}$ were determined, aliphatic side chains resonances were assigned by analysis of the following spectra: HNHA, HNHB, [^{1}H, ^{13}C]-CT-HSQC, and two HCCH-TOCSY spectra. In principle, a [^{1}H, ^{13}C]-CT-HSQC and a HCCH-TOCSY spectrum will suffice in the process of side chain resonance assignment. However the redundancy of information generated by use of the five spectra lends a higher degree of confidence to any assignments made rather than using the [^{1}H, ^{13}C]-CT-HSQC and HCCH-TOCSY spectra alone.

The [^{1}H, ^{13}C]-CT-HSQC spectrum, which shows all of the aliphatic ^{1}H-^{13}C couplings in $StpA_{91-134}$, was used as a reference spectrum (see Section 3.2.3.1). The Hα and Hβ resonance assignments, identified using the HNHA and HNHB spectra (see Section 3.2.3.2), together with the Cα and Cβ resonance assignments taken from experiments such as CBCACONH or HNCACB, allowed most of the resonances in the [^{1}H, ^{13}C]-CT-HSQC to be assigned. Where appropriate, the Cγ, Cδ, Cε, Hγ, Hδ, and Hε nuclei resonances were assigned by comparing the intra-residue correlations shown by the ^{15}N-TOWNY-HSQC spectrum, HCCONH spectrum, CCONH spectrum and the two HCCH-TOCSY spectra (shown in Section 3.2.3.3). In order to assign the side chain resonances of tryptophan, phenylanaline and tyrosine, a [^{1}H, ^{13}C]-CT-HSQC spectrum and a HCCH-TOCSY spectrum with modified transmitter offsets and a CBHD spectrum were acquired.

Some residues contain methylene (-CH_2-) groups; for example, the Cβ nuclei of residues such as serine, phenylanaline, arginine, lysine and tryptophan are methylene groups. In these cases one of either two phenomena can occur. When the side chain is rigid and spatially defined, for example, in cases where that side chain is involved in making inter-residue contacts, these two ^{1}H nuclei can occupy different spatial and chemical environments, and thus can have different chemical shifts. In other words, these two ^{1}H nuclei have stereospecifically distinct resonances. On the other hand, when the side chain is very flexible and structurally undefined, for example, when that side chain is pointing out into solution from the surface of the protein, the two ^{1}H nuclei are virtually indistinguishable. In this case the two ^{1}H nuclei have identical chemical shifts, and are said to be *degenerate*.

The residues alanine, leucine, isoleucine, valine, methionine and threonine have methyl groups (-CH_3) on their side chains. These methyl groups rotate rapidly, thus the three ^1H nuclei are structurally indistinguishable, resulting in one strong, degenerate signal corresponding to all three ^1H nuclei.

3.2.3.1 [^1H, ^{13}C]-CT-HSQC

The [^1H, ^{13}C]-CT-HSQC spectrum displays the correlations between a ^1H nucleus and the ^{13}C nucleus to which it is attached (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991; Vuister and Bax, 1992). The [^1H, ^{13}C]-CT-HSQC of $StpA_{91-134}$ is shown in Figure 3.9, and is significantly more complex than the [^1H, ^{15}N]-HSQC due to the presence of all of the ^1H-^{13}C resonances of the aliphatic side chains; hence the resonance assignments are not shown.

3.2.3.2 HNHA and HNHB

The HNHA spectrum shows the correlations between the amide HN and the Hα nucleus of the same residue (Kuboniwa *et al.*, 1994; Vuister and Bax, 1993). The signals corresponding to the Hα and the amide HN are opposite in phase relative to each other. Some examples of 2D slices of the HNHA spectrum showing resonances corresponding to glutamine-26, glycine-40 and serine-42 are shown in Figure 3.10. This spectrum can be used to distinguish glycine, which uniquely contains two Hα nuclei. The HNHB spectrum displays the correlations between the amide HN and the Hβ nuclei of the same residue (Archer *et al.*, 1991; Bax *et al.*, 1994). Some sample 2D slices of the HNHB spectrum (showing resonances corresponding to residues threonine-22, glutamine-26 and isoleucine-48) are shown in Figure 3.11.

3.2.3.3 HCCH-TOCSY

The HCCH-TOCSY spectrum displays resonances from any given diagonal cross peak in the HCCH-TOCSY to all ^1H nuclei found on that residue via ^1H-^{13}C-^{13}C-^1H scalar connectivities (Kay *et al.*, 1993). The diagonal cross peaks of the HCCH-TOCSY spectrum correspond to cross peaks found in the [^1H, ^{13}C]-CT-HSQC.

Two HCCH-TOCSY spectra of $StpA_{91-134}$ were acquired, with mixing times τ_m of 6 milliseconds (ms) and 21 ms. 2D slices of the two overlaid HCCH-TOCSY spectra are shown in Figure 3.12, with contours coloured green and purple respectively. The

experimental variable τ_m determines for how long the magnetisation is allowed to transfer between nuclei; the number of bonds along the aliphatic side chains, over which the magnetisation is transferred, can be changed by adjusting this value. Hence the HCCH-TOCSY experiment with the short τ_m of 6 ms shows correlations between ^{1}H nuclei separated by one or two ^{13}C-^{13}C bonds, whereas the HCCH-TOCSY experiment with the τ_m of 21 ms shows correlations between all intra-residue ^{1}H nuclei. Comparison of these two spectra allows one to distinguish ^{1}H and ^{13}C resonances that derive from adjacent ^{13}C-^{13}C nuclei, and scalar couplings between nuclei further apart in the side chain. This approach is particularly useful for assigning long side chains of residues such as arginine, proline, lysine, leucine and isoleucine. For example, as shown in Figure 3.12, the Hα, Hβ, Hγ, Hδ and Hε resonances correlated with lysine-31 ^{13}Cα are visible in the HCCH-TOCSY (with τ_m of 21 ms), whereas the HCCH-TOCSY (with τ_m of 6 ms) shows a strong diagonal cross peak, a cross peak of medium intensity corresponding to the Hβ nuclei, and weak cross peaks corresponding to the Hγ nuclei. There are no cross peaks corresponding to the Hδ and Hε nuclei in the HCCH-TOCSY (with τ_m of 6 ms).

3.2.3.4 Aromatic side-chain resonance assignments

Due to the large differences between the ranges of chemical shift seen in aliphatic and aromatic carbon nuclei, it was necessary to acquire a [^{1}H, ^{13}C]-CT-HSQC and HCCH TOCSY spectrum of $StpA_{91-134}$ optimised to detect aromatic side chain resonances. The aromatic [^{1}H, ^{13}C]-CT-HSQC shows correlations between ^{1}H nuclei and ^{13}C nuclei of the side chains of phenylanaline, tryptophan and tyrosine (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991; Vuister and Bax, 1992). Figure 3.13 shows a [^{1}H, ^{13}C]-CT-HSQC of the aromatic side chains resonances. The aromatic HCCH-TOCSY spectrum displays correlations between ^{1}H nuclei and ^{13}C nuclei found in aromatic rings (Kay *et al.*, 1993). Figure 3.14 shows the aromatic HCCH-TOCSY spectrum superimposed upon the aromatic ^{13}C-NOESY-HSQC spectrum shown in red and blue, respectively (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991). The CBHD [also referred to as (HB)CB(CGCD)HD] spectrum shows correlations between the Cβ and the Hδ nuclei on the of an aromatic side chain (Yamazaki *et al.*, 1993). This 2D spectrum allowed any each aromatic spin system to be corresponded to specific residues in $StpA_{91-134}$ (data not shown).

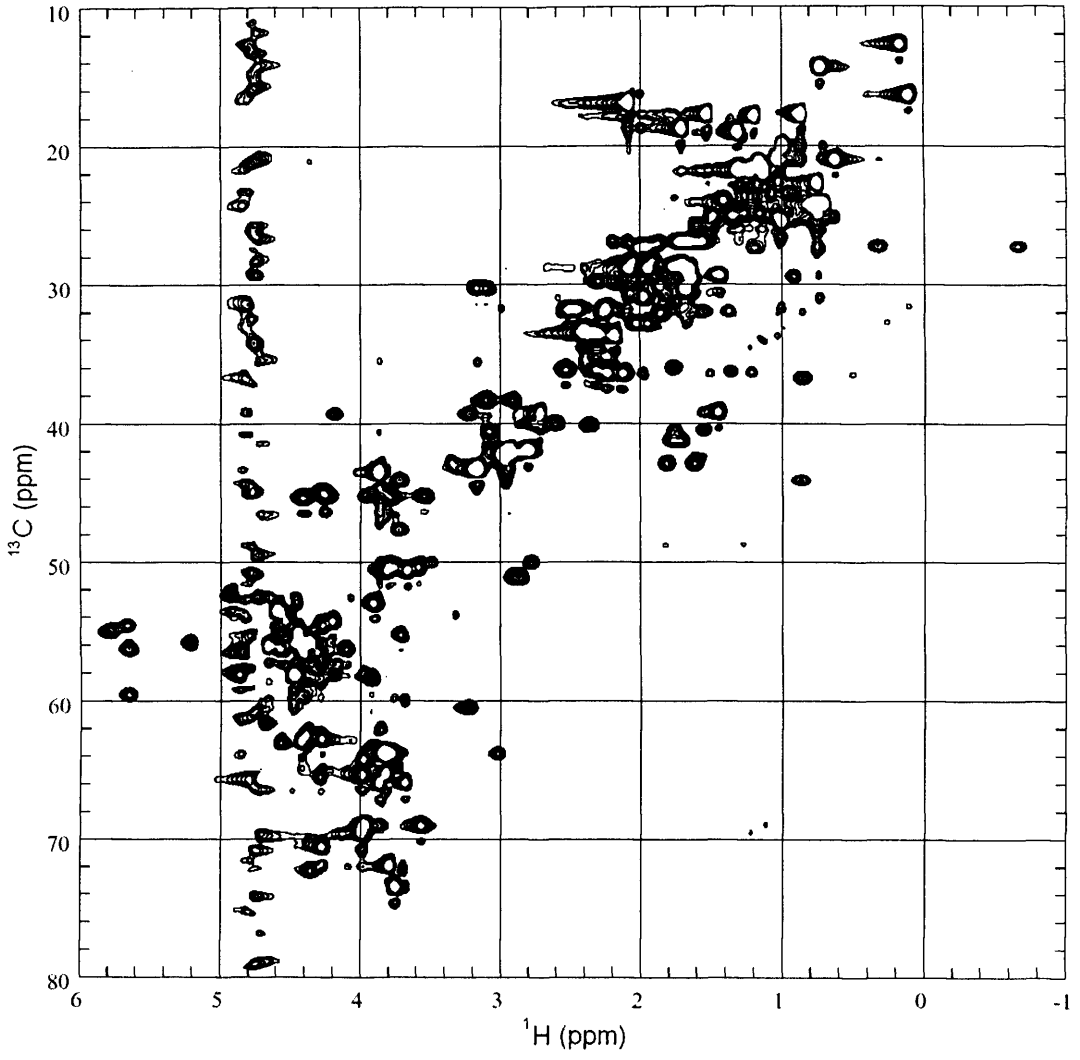


Figure 3.9
[^{1}H, ^{13}C]-CT-HSQC spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ showing aliphatic resonances only, recorded on a 500 MHz Varian spectrometer at 25°C. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3.

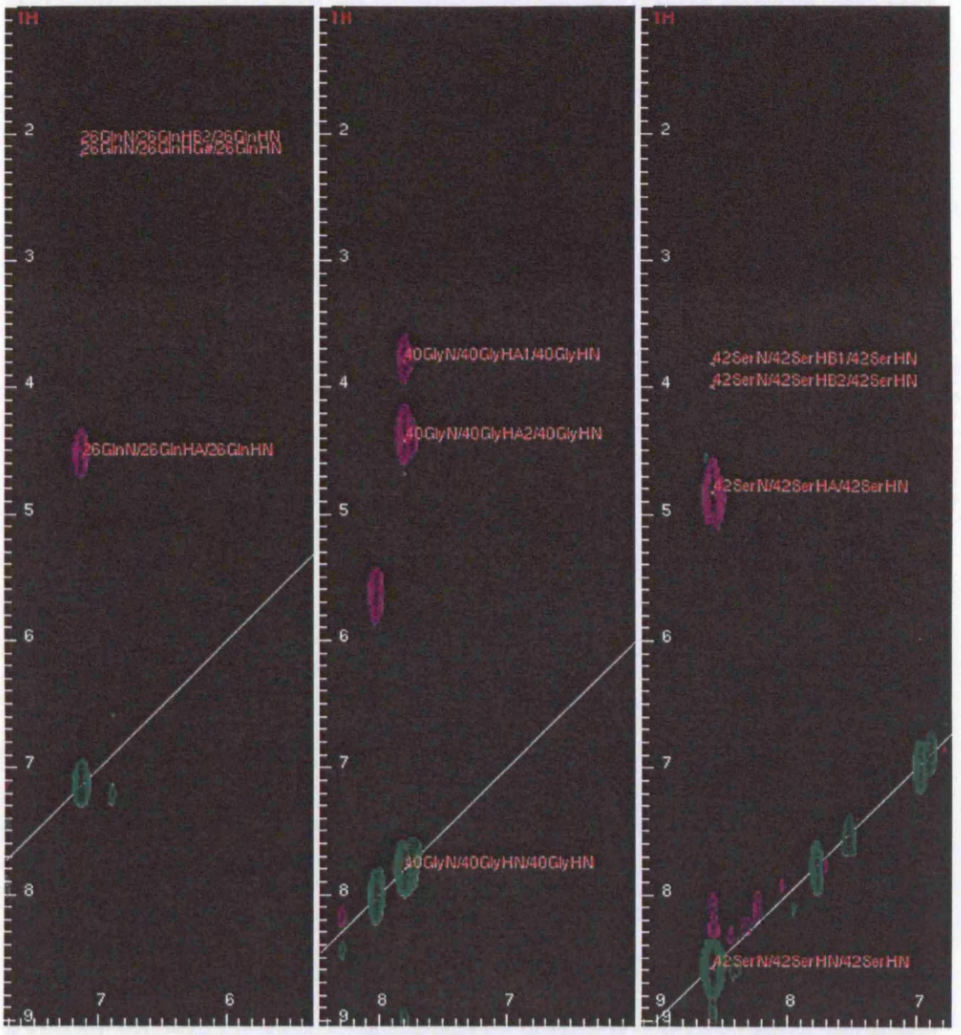


Figure 3.10

A selection of 2D slices of the HNHA spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$, showing resonances corresponding to residues Gln-26, Gly-40 and Ser-42 of StpA$_{91\text{-}134}$. Diagonal NH cross peak contours are coloured green, whereas Hα resonance cross peaks are coloured purple, due to the different phases of the two types of cross peaks. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3. Both spectra were recorded on a 500 MHz Varian spectrometer at 25°C.

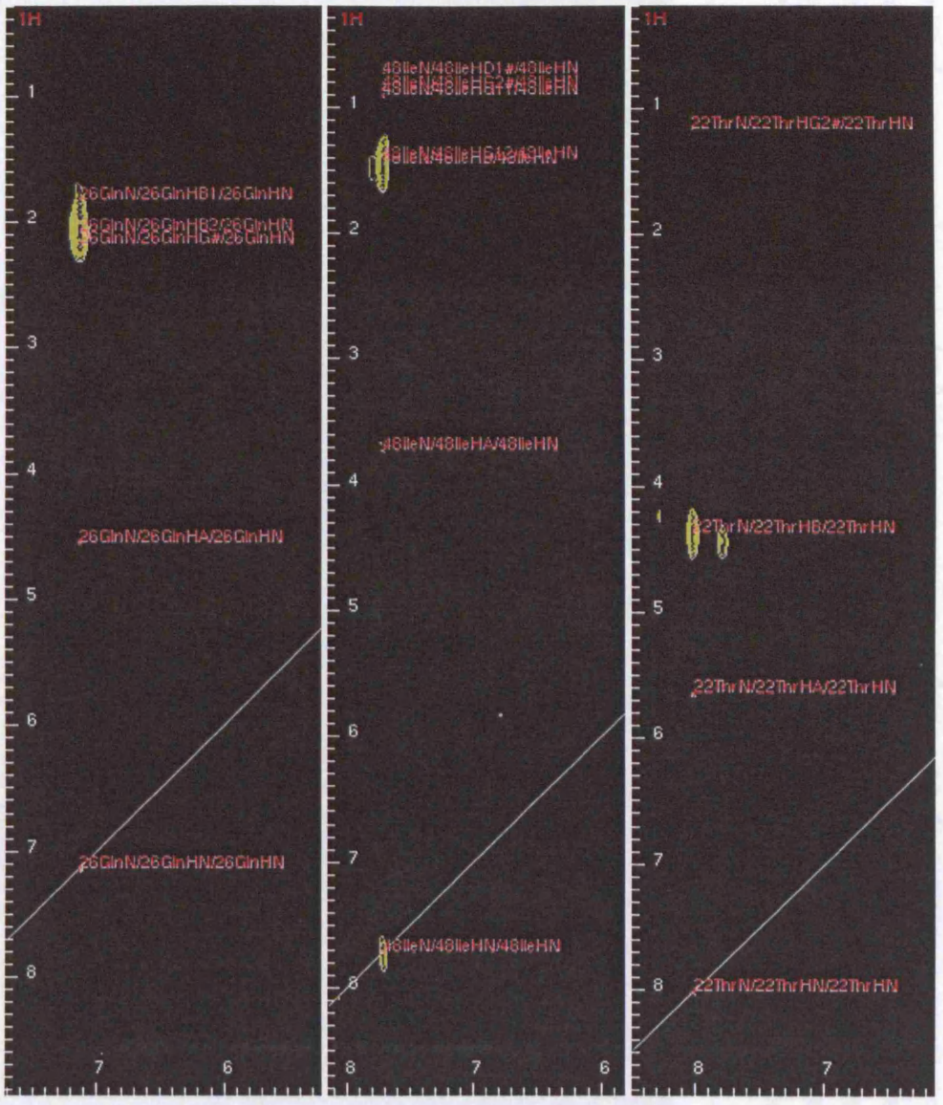


Figure 3.11

Three sample 2D slices of the HNHB spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$, showing resonances corresponding to residues 26, 48 and 22 of the amino acid sequence of $StpA_{91-134}$. The HNHB spectrum chows correlations between the backbone amide NH nuclei with the Hβ nuclei of the same residue (with contours coloured yellow). The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3. The HNHB spectrum was recorded on a 500 MHz Varian spectrometer at 25°C.

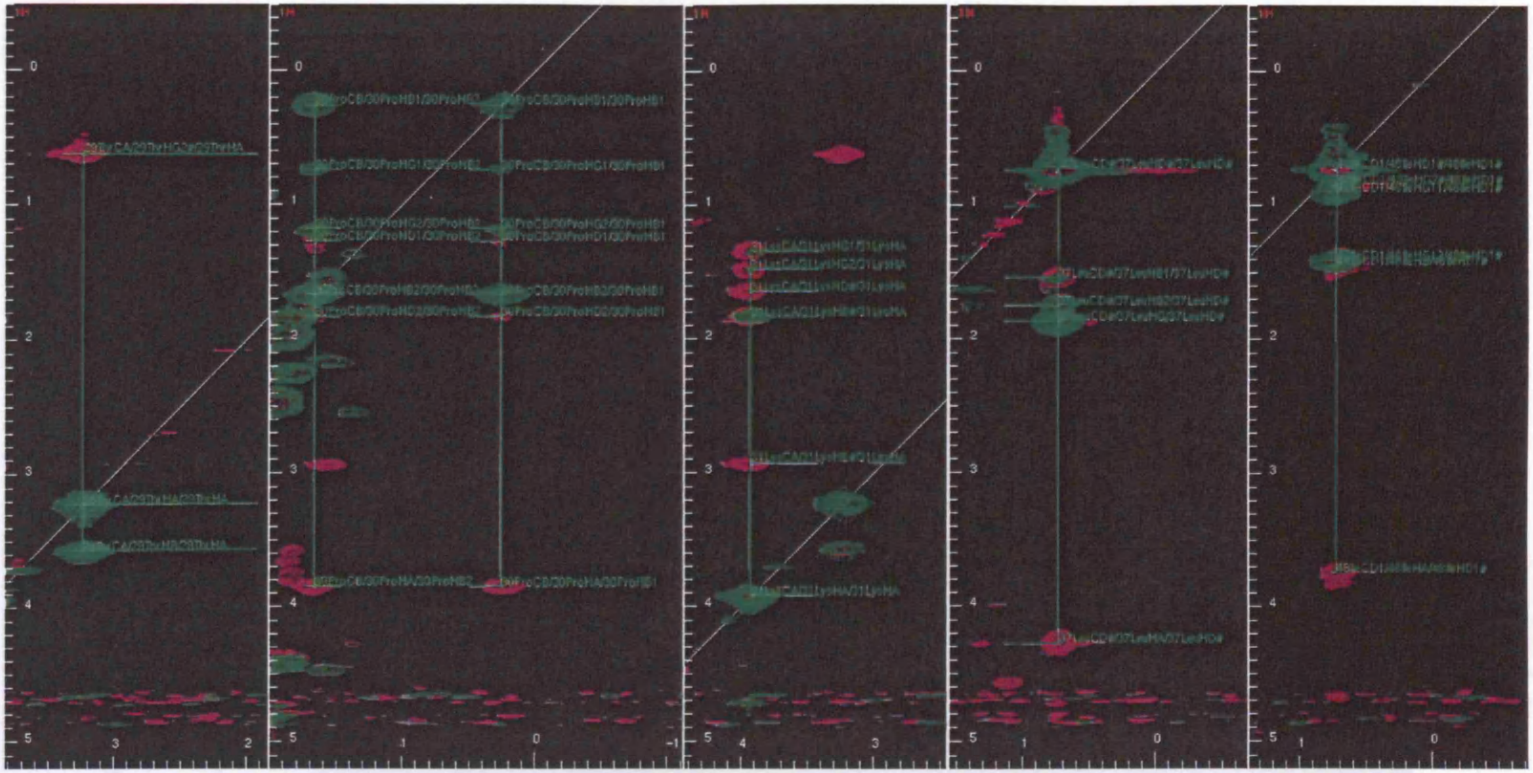


Figure 3.12

Sample 2D slices of the two HCCH-TOCSY spectra of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$: the HCCH-TOCSY with τ_m of 6 ms, with contours coloured green, is superimposed upon the HCCH-TOCSY with τ_m of 21 ms, with contours coloured purple. The 2D slices correspond to the following diagonal resonances: Thr-29 Cα, Pro-30 Cβ, Lys-31 Cα, Leu-37 Cδ#, and Ille-48 Cδ1. (Leu-37 of StpA$_{91\text{-}134}$ has two Cδ methyl groups that are structurally and chemically indistinguishable, giving rise to one cross peak. Degenerate nuclei, such as the two Cδ methyl group of Leu-37, are annotated with a '#'.) The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3. The HCCH-TOCSY spectra were recorded on a 500 MHz Varian spectrometer at 25°C.

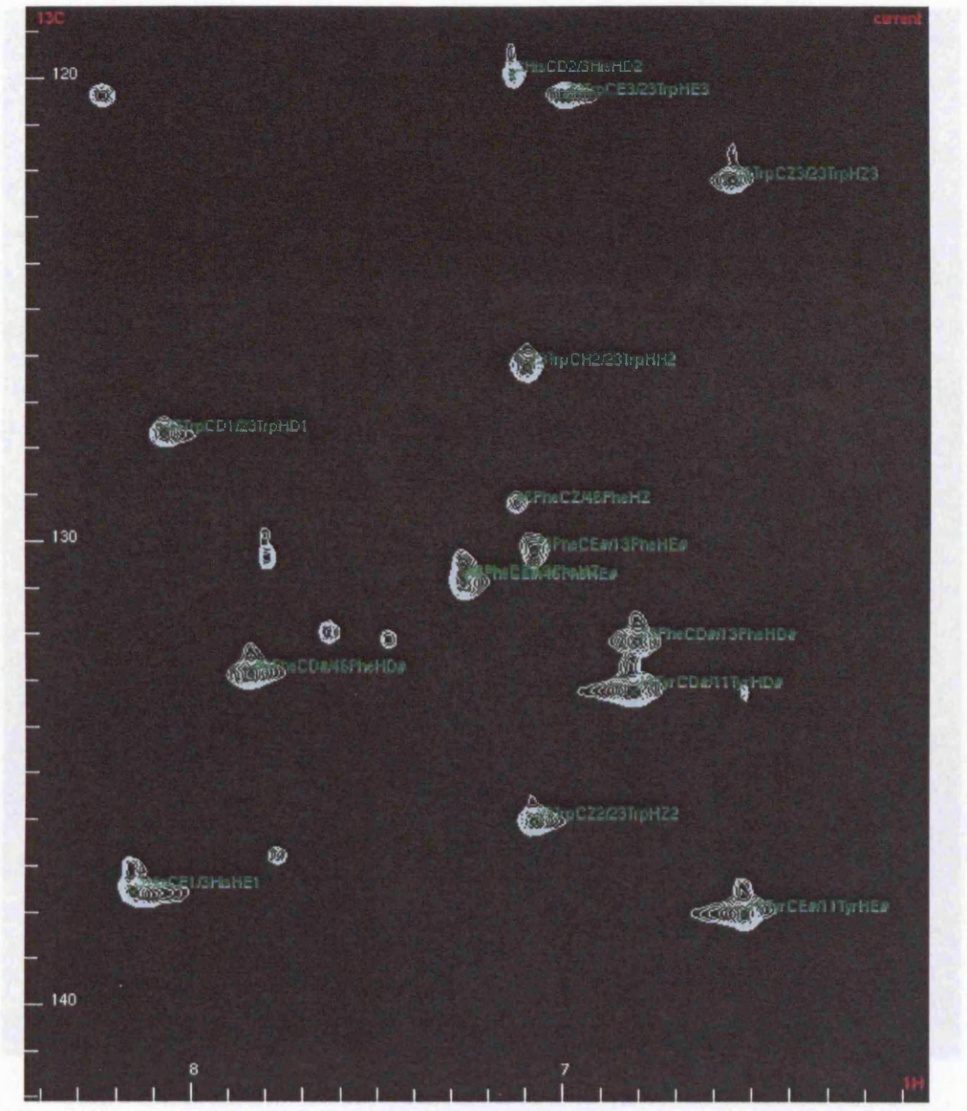


Figure 3.13
A [^{1}H, ^{13}C]-CT-HSQC of the aromatic side chains of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$, recorded on a 500 MHz Varian Spectrometer at 25°C. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$ with 10 mM sodium phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3.

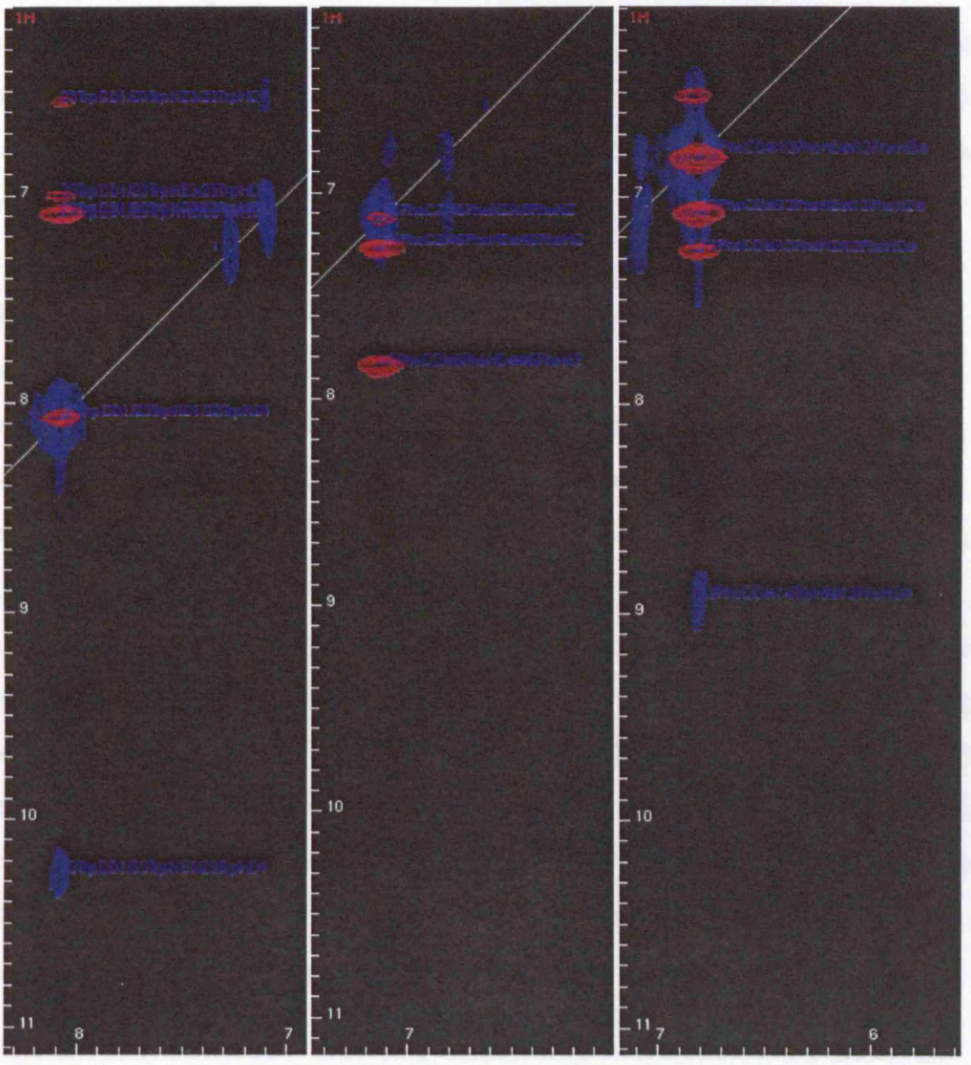


Figure 3.14
HCCH TOCSY and ^{13}C-NOESY-HSQC spectra (with contours coloured red and blue, respectively) showing the aromatic side chains of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$, recorded on a 500 MHz Varian Spectrometer at 25°C. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM sodium phosphate pH 6.5, 10 % D2O (v/v), 100 mM NaCl, 100 μM EDTA and 10 μM NaN_3.

3.2.4 Secondary structure determinations

Information concerning the secondary structure of a protein can be deduced directly from NMR chemical shifts. The formation of secondary structure induces general and predictable changes in chemical shift regardless of amino acid type. By comparing the chemical shift of the Hα, Cα, and Cβ nuclei resonances with chemical shift values obtained from random-coils, it is possible to tentatively identify regions of helical or sheet secondary structure in a protein. Loop and turn structures may also change the chemical shifts compared to random coil values, but their effect on chemical shift is less predictable (Spera and Bax, 1991; Wishart *et al.*, 1991).

The chemical shift values of Hα, Cα, and Cβ resonances derived from random coils were subtracted from the appropriate chemical shifts values of Hα, Cα, and Cβ resonances of $StpA_{91-134}$. These secondary chemical shifts of Hα, Cα, and Cβ nuclei resonances are plotted with respect to the $StpA_{91-134}$ amino acid sequence in Figure 3.15, labelled as 'Δ^1Hα (ppm)', 'Δ^{13}Cα (ppm)' and 'Δ^{13}Cβ (ppm)', respectively. Negative Δ^{13}Cα, positive Δ^{13}Cβ and positive Δ^1Hα values suggest helix formation, whilst positive Δ^{13}Cα, negative Δ^{13}Cβ and negative Δ^1Hα values suggest β-strand formation. The secondary chemical shift values of the Hα, Cα, or Cβ resonances suggests $StpA_{91-134}$ has 2 β-strands (between Tyr-11 to Asp-15, and between Lys-21 to Trp-23) and two helices (between Lys-31 to Ala-38, and between Leu-43 to Asp-45). The secondary chemical shifts of the backbone carbonyl ('Δ^{13}CO') and backbone amide ^{15}N ('Δ^{15}N') are also shown in Figure 3.15; however there is little correlation between these values with $StpA_{91-134}$ secondary structure.

3.2.4.1 TALOS

A more rigorous approach to estimating secondary structure from chemical shift data has been described (Cornilescu *et al.*, 1999). A program named TALOS (or Torsion Angle Likelihood Obtained from Shift and sequence similarity) was used to compare chemical shift values of Hα, Cα, Cβ, CO and N of $StpA_{91-134}$ with a database of empirical values derived from known structures in order to make quantitative predictions of φ (phi) and ψ (psi) backbone torsion angles of $StpA_{91-134}$ (see Section 2.13.2) .

The data sets labelled 'φ' and 'ψ' in Figure 3.15 shows the predicted values and errors of the backbone dihedral angles, plotted against residue identity. Where all ten pairs

of backbone dihedral angles of the best database matches lay within the same region of the Ramachandaran diagram (corresponding to a prediction of high confidence), the data point was coloured green. Where only nine or less of the ten pairs of backbone dihedral angles lay within the same region of the Ramachandaran diagram, indicating either an ambiguous or poor prediction, the data point was coloured red. Only good predictions were used as restraints in the structural calculations of $StpA_{91-134}$ (described in Sections 3.2.6 and 3.2.7). The TALOS predictions (of high confidence) of φ and ψ backbone dihedral angles shown in Figure 3.15 indicate that there are four regions of secondary structure: two β-strands (between residues Tyr-11 to Thr-14 and Glu-19 to Thr-24) and two α-helices (between residues Pro-32 to Glu-39 and Leu-43 to Asp-45). These predictions of $StpA_{91-134}$ secondary structure agree with the secondary structure deduced from secondary chemical shift values of Hα, Cα, or Cβ resonances, and with the secondary structure of the C-terminus of the homologous protein $H\text{-}NS_{90-136}$ (deposited in the Protein Data Bank, with the identification code 1HNR) (Shindo *et al.*, 1995).

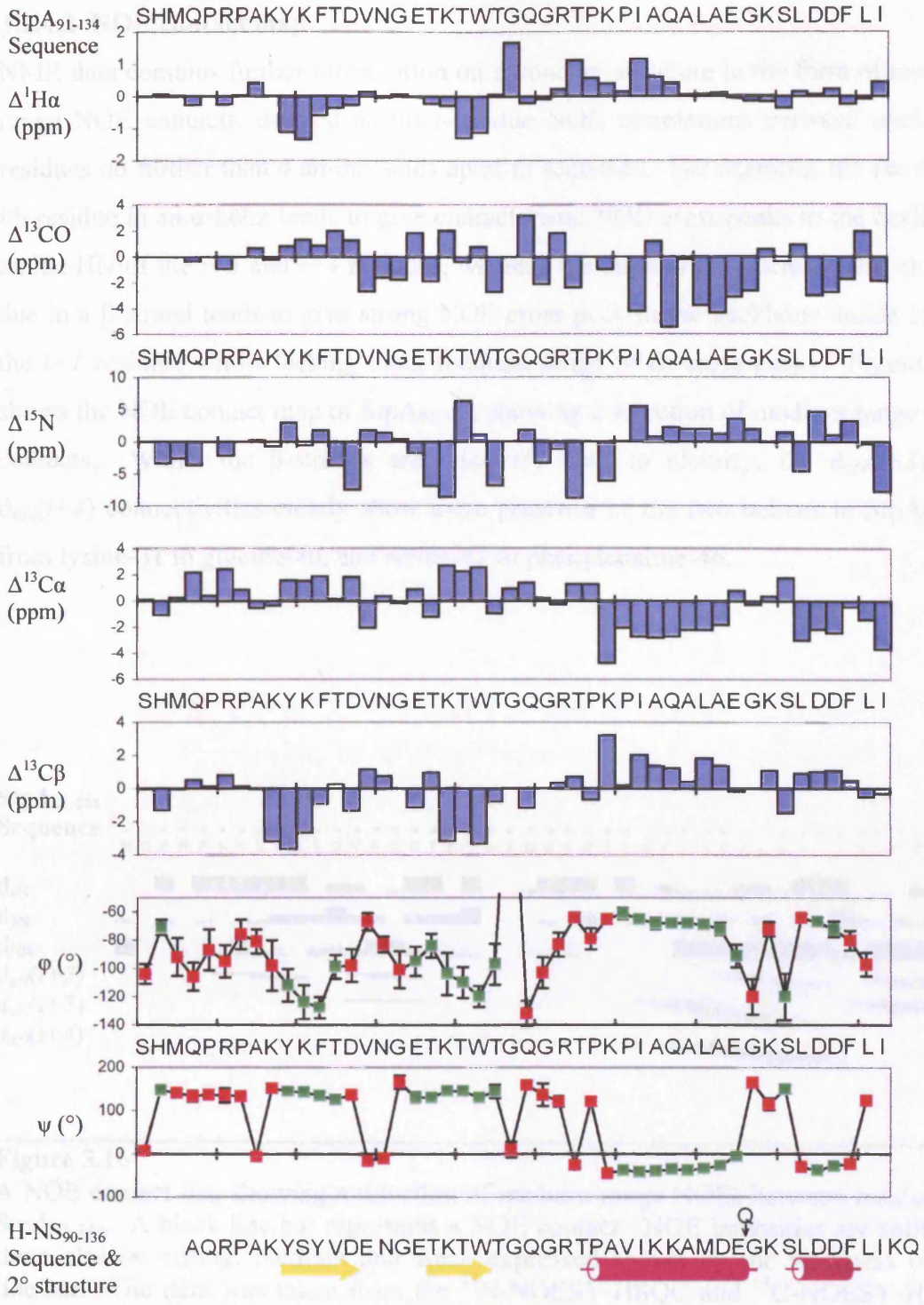


Figure 3.15

Summary of secondary structure of $StpA_{91-134}$ determined using NMR data. The rows labelled $\Delta^1H\alpha$, $\Delta^{13}CO$, $\Delta^{15}N$, $\Delta^{13}C\alpha$ and $\Delta^{13}C\beta$ show the secondary chemical shifts of the respective nuclei types, in ppm, plotted against $StpA_{91-134}$ amino acid sequence. TALOS φ and ψ backbone dihedral angles have also been plotted. Dihedral angles shown as a green and red dots represent TALOS predictions of high and low confidence, respectively. The amino acid sequence and secondary structure of the homologous H-NS_{90-136} structure, aligned with respect to $StpA_{91-134}$, is shown.

3.2.4.2 NOE contact map

NMR data contains further information on secondary structure in the form of medium range NOE contacts, defined as inter-residue NOE correlations between nuclei on residues no further than 4 amino acids apart in sequence. For example, the Hα of the *i*th residue in an α-helix tends to give characteristic NOE cross peaks to the backbone amide HN of the *i+3* and *i+4* residues, whereas the Hα and Hβ nuclei of the *i*th residue in a β-strand tends to give strong NOE cross peak to the backbone amide HN of the *i+1* residue, whilst lacking other medium range NOE cross peaks. Figure 3.16 shows the NOE contact map of $StpA_{91-134}$, showing a selection of medium range NOE contacts. Whilst the β-strands are relatively hard to identify, the $d_{\alpha N}(i+3)$ and $d_{\alpha N}(i+4)$ connectivities clearly show there presence of the two helices in $StpA_{91-134}$, from lysine-31 to glycine-40, and serine-42 to phenylanaline-46.

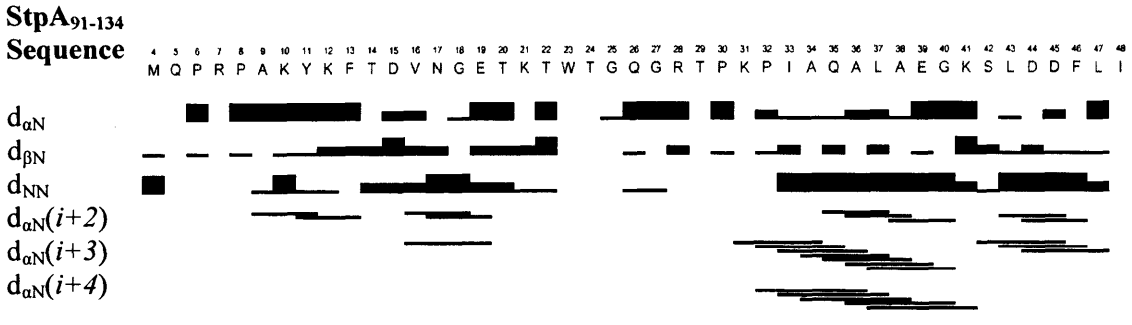


Figure 3.16
A NOE contact list, showing a selection of medium-range NOEs between residues of $StpA_{91-134}$. A black line/bar represents a NOE contact. NOE intensities are split into three classes: strong, medium and weak expressed shown by the thickness of the line/bar. The data was taken from the ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra of [^{1}H, ^{15}N, ^{13}C]-labelled $StpA_{91-134}$.

3.2.5 Relaxation analysis

The spin relaxation of the ^{15}N nucleus of the amide bond is determined by the dipolar coupling with the amide ^{1}H nucleus and chemical shift anisotropy (CSA) of the of the ^{15}N nucleus, where the CSA results in a small magnetic field generated as a result of the electron structure surrounding the ^{15}N. The relaxation characteristics of the ^{15}N nucleus can be correlated with the dynamic motion of the backbone amide HN vector, and hence the dynamic behaviour of a protein (Szyperski *et al.*, 1993).

Dynamic properties of $StpA_{91-134}$ were investigated by measuring relaxation rates of the nuclear magnetisation of the backbone amide HN. Residue-specific longitudinal relaxation time constants (T_1) were determined from an 'inversion recovery' experiment, where the bulk magnetisation vector M_Z is rotated by a 180° pulse to give a $-M_Z$ bulk magnetisation vector (see Figure 3.2). The rate of decay of the vector $-M_Z$ back to thermal equilibrium, in other words recovery of the vector M_Z, is related to T_1. The bulk magnetisation magnitude is measured by application of a 90° RF pulse after an incremented delay to rotate the longitudinal magnetisation into the transverse plane (Kay *et al.*, 1989; Kay *et al.*, 1992b; Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991). The residue-specific T_1 relaxation constants are shown in Figure 3.17.

The residue-specific transverse relaxation time constant (T_2) was determined from a 'spin-echo' pulse sequence, where the bulk magnetisation vector M_Z is subjected to a 90° RF pulse, followed by a 180° RF pulse after a variable delay *t*. The 90 ° RF pulse generates transverse magnetisation, which gradually loses phase coherence during the delay according to T_2; different ^{15}N nuclei precess at slightly different rates, spreading the distribution of their phases in the transverse plane. The 180 ° RF pulse inverts the distribution of the phases of the ^{15}N nuclei, such that after an equal delay of *t*, the magnetisation is refocused, and phase coherence recovered. The residue-specific T_2 relaxation constants are shown in Figure 3.17.

The residue specific [^{1}H, ^{15}N]-heteronuclear NOE magnitude describes the heteronuclear dipolar coupling between the ^{1}H and the ^{15}N nuclei in the amide HN group. The heteronuclear NOE provides a measure of the rate of the amide HN bond vector rotation. The ^{1}H nucleus of the amide group is excited with a RF pulse; the ^{15}N nuclear spin is then affected according to the heteronuclear dipolar coupling. The [^{1}H, ^{15}N]

heteronuclear NOE is calculated as a ratio of the signal intensity of the ^{15}N nucleus of the amide HN with or without ^{1}H nuclear spin saturation (Kay *et al.*, 1989; Kay *et al.*, 1992b; Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991). The residue-specific [^{1}H, ^{15}N] heteronuclear NOE values are shown in Figure 3.17.

The residue specific T_1/T_2 ratios and the [^{1}H, ^{15}N] heteronuclear NOEs give an indication of the flexibility of the backbone of a protein. Low T_1/T_2 ratios qualitatively suggest higher levels of flexibility, implying that the polypeptide backbone encompassing the residues Lys-21 to Trp-23 and Ile-33 are particularly flexible (Figure 3.17). In contrast, according to the secondary chemical shift analysis and the TALOS predictions (discussed in Section 3.2.4), these residues are involved in secondary structure. The residues suggested by the secondary structure determinations to be involved in loops, which are anticipated to show a comparatively high level of flexibility, however do not show a low T_1/T_2 ratio. [^{1}H, ^{15}N] heteronuclear NOE values (whilst dependent on magnetic field strength B_0) can vary from 0.8 (implying high rigidity of the backbone), to around 0.5 (for flexible loops and termini) to lower or negative values (for disordered regions) (Kay *et al.*, 1989). The [^{1}H, ^{15}N] heteronuclear NOE values for $StpA_{91\text{-}134}$ suggest residues Gln-5 to Ala-9 are disordered: the rest of the protein, with [^{1}H, ^{15}N] heteronuclear NOE values ranging between 0.5 and 0.7, is flexible. There does not seem to be a strong relationship between secondary structure and [^{1}H, ^{15}N] heteronuclear NOEs. This inherent high level of flexibility exhibited by $StpA_{91\text{-}134}$ may be a consequence of the small size of the domain, resulting in a relatively loose structure.

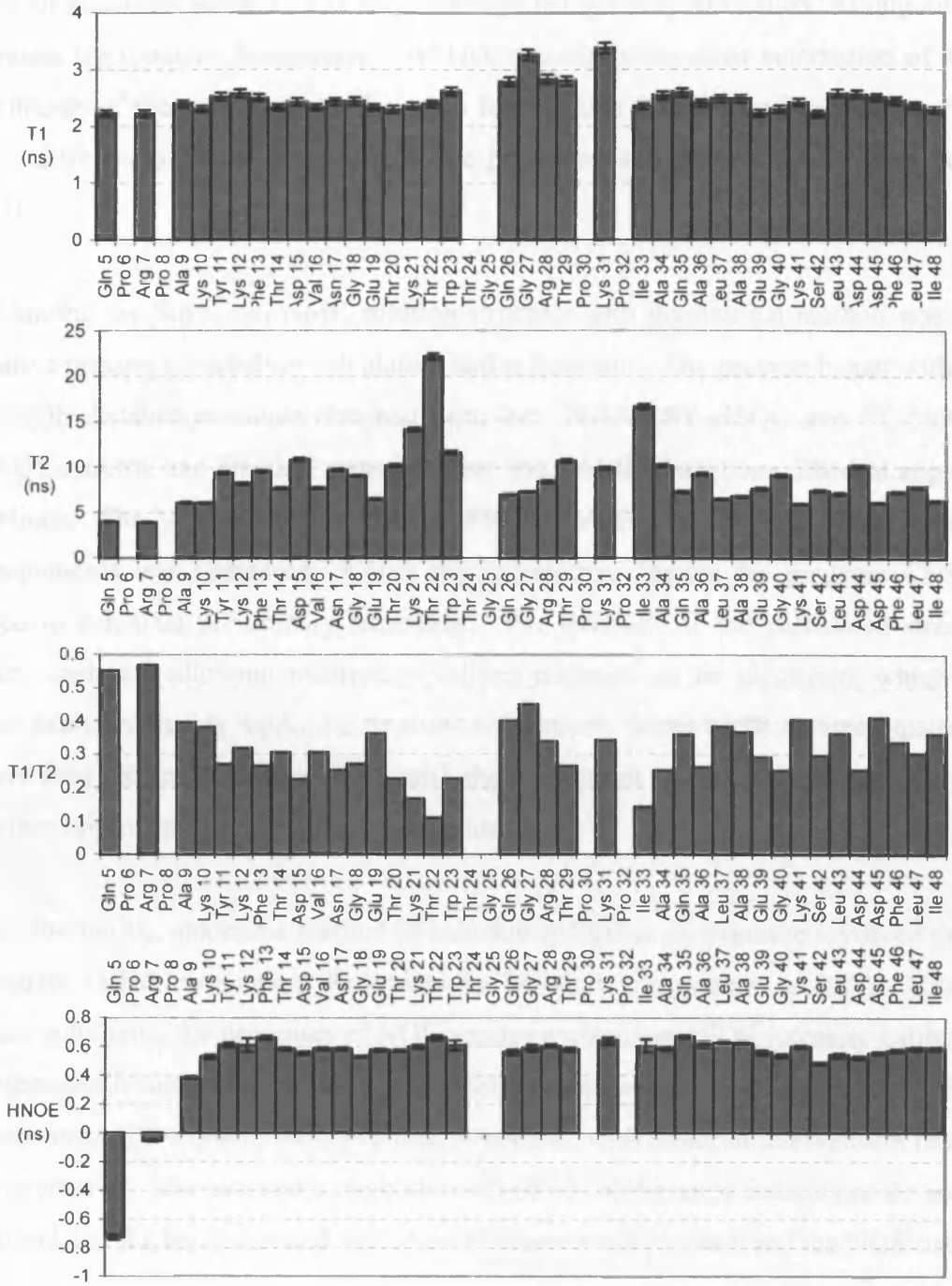


Figure 3.17
Residue-specific ^{15}N relaxation data of $StpA_{91-134}$. The values of T_1, T_2, the ratio of T_1/T_2, and [^1H, ^{15}N] heteronuclear NOEs are plotted with respect to the amino acid sequence of $StpA_{91-134}$.

3.2.6 Structure calculation - overview

Two methods were used to determine the structure of $StpA_{91-134}$: a *manual method* and an *automatic method*. The latter method, using a program called 'Ambiguous Restraints for Iterative Assignment' (ARIA), essentially involves automation of all the protocols of the manual method, with a few notable differences (see Section 3.2.7). The differences of the two methods are presented and discussed later (see Section 3.3).

Obtaining the $StpA_{91-134}$ NMR solution structure with the manual method was an iterative process of structure calculation and refinement. The process began with a list of NOE distance restraints obtained from the ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra, and dihedral restraints from the TALOS backbone dihedral angle predictions. The NOE distance restraints were calibrated and the lack of stereospecific assignments was addressed. Using the restraints as inputs, the program CNS was used to calculate preliminary structures. The qualities of the calculated structures were analysed, allowing incorrectly defined restraints to be identified, which were then excluded during $StpA_{91-134}$ structure refinement. More NMR structure ensembles were then recalculated using CNS with the new list of restraints, until there were no further improvements in $StpA_{91-134}$ structure.

The alternative, automatic method of calculating $StpA_{91-134}$ structure involved using a program called 'Ambiguous Restraints for Iterative Assignment' (ARIA). This program automated the processes of NOE spectra assignment, NOE intensity calibration, stereospecific assignments, $StpA_{91-134}$ structure calculation and refinement, and finally assessment of the quality of the refined structures, with minimal intervention from the program user. The automatic method involved use of the same datasets as the manual method, i.e. the backbone and side chain resonance assignments and the NOE data.

3.2.6.1 Distance restraints: the ^{13}C- and ^{15}N-NOESY-HSQC spectra

The nuclear Overhauser effect (NOE) in experiments such as NOESY-HSQC spectra results in correlations between pairs of nuclei in close proximity, and occurs with detectable intensity between nuclei that are less than 6 Å apart. The magnitude of NOEs between nuclei is proportional to r^{-6} (where r is distance between the nuclei). The ^{15}N-NOESY-HSQC shows correlations between backbone or side-chain amide HN

and all ^{1}H nuclei within 6 Å (shown in Figure 3.18), and the ^{13}C-NOESY-HSQC shows correlations between any ^{1}H nuclei within 6 Å (shown in Figure 3.19) (Kay *et al.*, 1992a; Palmer A.G. *et al.*, 1991). Solution structures are calculated by determining the spatial orientation of the backbone and side chains that satisfies the thousands of distance restraints and backbone dihedral restraints derived from the NOE data, using a program called CNS (Brunger *et al.*, 1998).

Short-range NOEs are defined as those specifying intra-residue distances; medium-range NOEs specify distances between residues separated by 4 amino acids or less in the amino acid sequence; long-range NOEs specify distances between residues separated by more than 4 amino acids. The majority of NOE correlations observed with NOESY-HSQC spectra of $StpA_{91\text{-}134}$ were either short or medium range. Short-range NOEs were assigned by comparing the NOESY-HSQC spectra to those that only show intra-residue correlations such as HCCH-TOCSY and ^{15}N-TOWNY-HSQC. Medium- and long-range NOEs were assigned during the process of structural refinement of $StpA_{91\text{-}134}$, described in Section 3.2.6.4. Alternatively, the automatic method described in Section 3.2.7 also resulted in assignment of the $StpA_{91\text{-}134}$ NOESY-HSQC spectra.

3.2.6.2 NOE distance calibration

The nuclear Overhauser effect between two nuclei is proportional to the inverse sixth power of the distance between those nuclei. The NOE decays so rapidly with distance that the NOE is usually undetectable between nuclei more 6 Å apart. The relationship between NOE cross-peak intensity and the distance between two nuclei was calibrated by looking at pairs of nuclei separated by known interatomic distances in the structure; for example, the empirically determined distances between backbone nuclei in a α-helix are well documented (Roberts, 1993). However, rather than assigning an interatomic distance for every NOE cross-peak intensity observed, the NOE cross-peak intensities were split into three categories; strong, medium and weak intensities, simplifying the calculation of $StpA_{91\text{-}134}$ structure.

Table 3.1 contains known ^{1}H-^{1}H distances that were compared with NOE cross-peak intensities corresponding to equivalent distances in $StpA_{91\text{-}134}$ NOESY-HSQC spectra. The natural logarithm of the NOE cross-peak intensity was plotted against distance

for both ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra (Figure 3.20). Using the NOE intensity calibration the three categories of strong, medium and weak intensities were defined (corresponding to the distances <2.9Å, <3.8Å and <5.5Å, respectively) for the ^{15}N-NOESY-HSQC spectrum and two categories of strong and weak intensities were defined (corresponding to the distances <2.5Å and <5.5Å, respectively) for the ^{13}C-NOESY-HSQC spectrum. All NOE cross-peak intensity data were then categorised according to their intensities.

Distance (Å)	α-helix	β-strand
$d_{\alpha N}$	3.5	2.2
$d_{\alpha N}(i+2)$	4.4	-
$d_{\alpha N}(i+3)$	3.4	
$d_{\alpha N}(i+4)$	4.2	-
d_{NN}	2.8	4.3
$d_{NN}(i+2)$	4.2	-

Table 3.1
Known interatomic distances in α-helices and β-strands

3.2.6.3 Stereospecificity

As discussed in Section 3.2.3, the methylene group can exhibit distinct ^{1}H chemical shifts if the two ^{1}H nuclei are in chemically distinct environments. This occurs frequently when the methylene group resides in a region of rigid structure. A program called Anglesearch, which calculates torsion angles based on data such as HNHA, HNHB, and NOESY-HSQC spectra cross-peak intensities, can be used to assign the stereospecific ^{1}H nuclei (Polshakov *et al.*, 1995). However attempts to use Anglesearch failed and will not be discussed further.

Without stereospecific assignments, NOE data from the ^{1}H nuclei from the methylene groups would not constrain the calculated structure correctly around the unassigned methylene group and cannot be used during the calculation of $StpA_{91-134}$ structure. Therefore these methylene groups were replaced by a pseudoatom placed between the two ^{1}H nuclei, and information derived from NOE intensities averaged. Methyl groups ($-CH_3$), the ^{1}H nuclei of which never have distinct chemical shifts, were also replaced with a pseudoatom for the purposes of calculating $StpA_{91-134}$ structure.

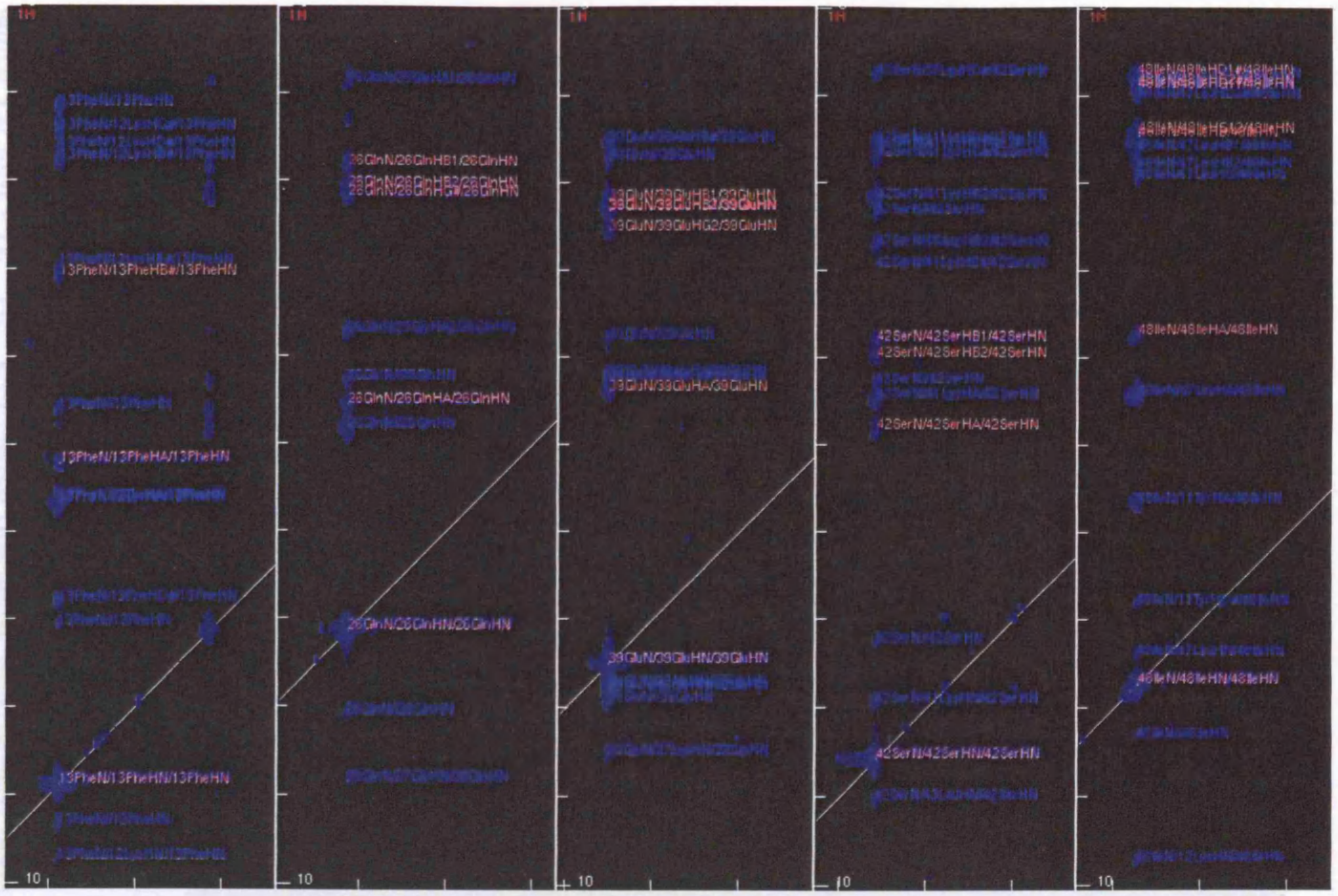


Figure 3.18
Sample 2D slices of the ^{15}N-NOESY-HSQC spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$, with contours and assignments shown in blue. Intra-residue resonance assignments are coloured red. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM Sodium Phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3. The HCCH-TOCSY spectra were recorded on a 500 MHz Varian spectrometer at 25°C.

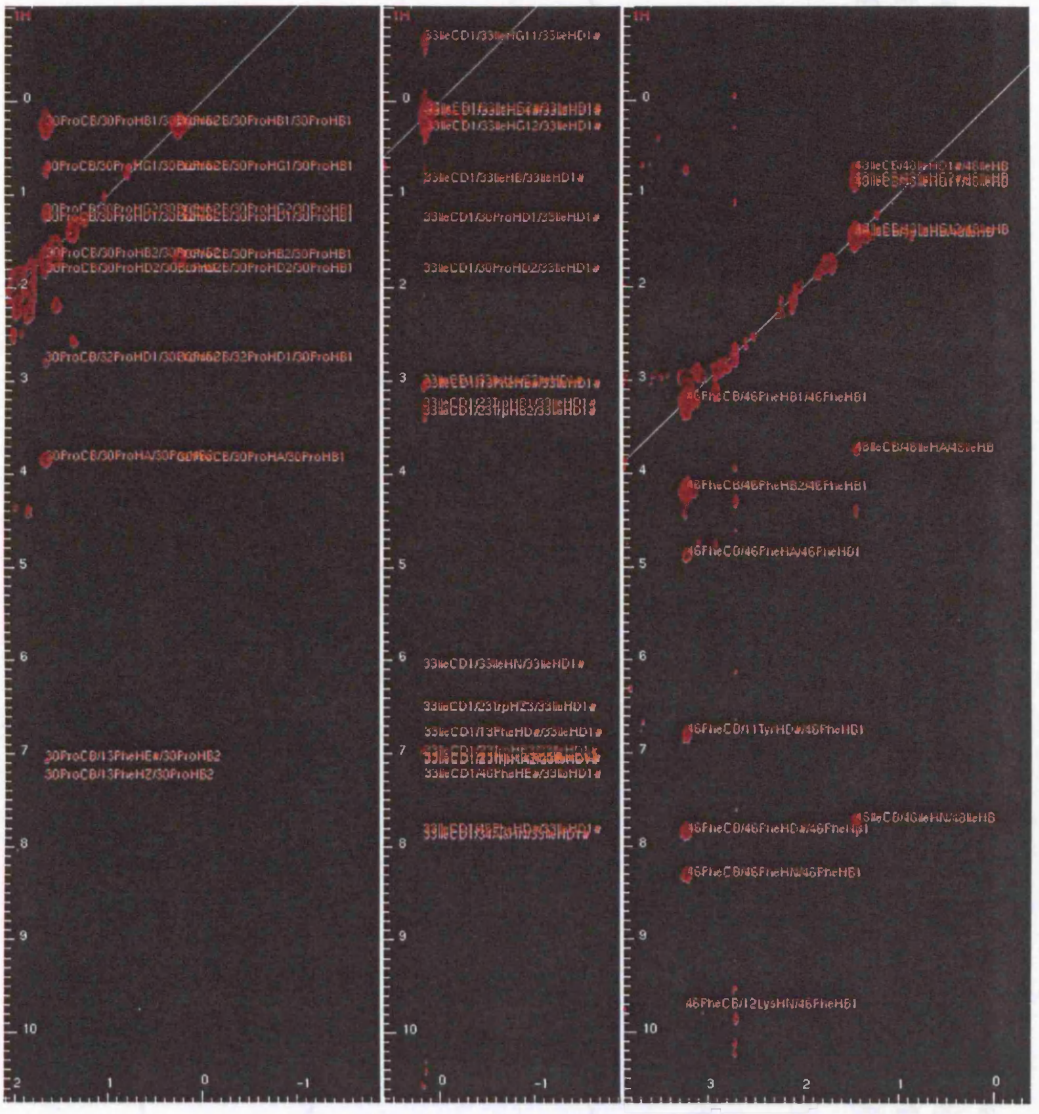


Figure 3.19
A selection of 2D slices of the ^{13}C-NOESY-HSQC spectrum of [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$, recorded on a 500 MHz Varian Spectrometer at 25°C. The sample contained 1.5 mM [^{1}H, ^{15}N, ^{13}C]-labelled StpA$_{91\text{-}134}$ with 10 mM Sodium Phosphate pH 6.5, 10 % D_2O (v/v), 100 mM NaCl, 100 µM EDTA and 10 µM NaN_3.

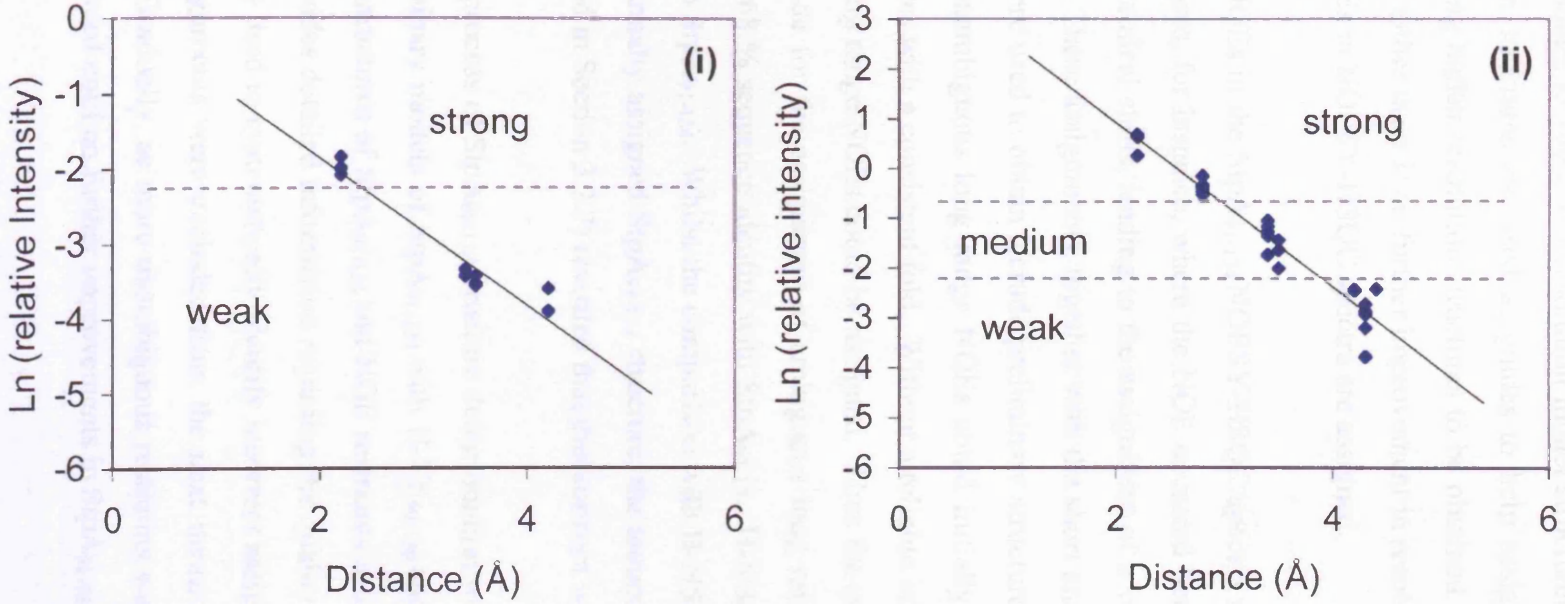


Figure 3.20

NOE cross-peak intensity calibration. The natural logarithm of intensities of NOE cross-peaks in the ^{15}N- and ^{13}C-edited NOESY-HSQC spectra that correspond to known distances were plotted against distance (Å). The line of best fit was used to calibrate the NOE intensities in the ^{15}N- and ^{13}C-edited NOESY-HSQC spectra. The categories of strong, medium and weak intensities, into which the relative intensities are divided after the calibration process, are also shown.

3.2.6.4 Structure refinement

In the initial stages of determining protein structure by NMR, a few long range NOEs are usually enough to allow low resolution model structures to be calculated. These low-resolution structures are used as guides to help assign ambiguous long range NOEs, allowing higher resolution structures to be obtained. This iterative process is repeated until either there is no further improvement in resolution of structure, or until all of the NOEs in NOESY-HSQC spectra are assigned.

Long range NOEs in the $StpA_{91-134}$ NOESY-HSQC spectra were only assigned in unambiguous cases, for instance, where the NOE occurred between two ^{1}H nuclei having unique chemical shifts, leading to the assignment of a very small number of long range NOEs. These assignments, together with the short and medium range NOE assignments, were used to obtain a crude preliminary structure of $StpA_{91-134}$. However, not enough unambiguous long range NOEs could initially be assigned to generate crude structures with a consistent fold. Without a reliable initial structure, none of the ambiguous long range NOEs could be assigned. Thus the structure of $H\text{-}NS_{90-136}$ was used as a guide for the assignment of ambiguous long range NOEs (Shindo *et al.*, 1995). With 68 % sequence identity with $StpA_{91-134}$, $H\text{-}NS_{90-136}$ was likely to have a similar fold to $StpA_{91-134}$. Whilst the comparison with $H\text{-}NS_{90-136}$ structure could have biased the manually assigned $StpA_{91-134}$ structure, the automatic analysis of the NMR data (described in Section 3.2.7) revealed that this concern was unfounded.

The iterative process of $StpA_{91-134}$ structure determination was carried out by comparing the preliminary models of $StpA_{91-134}$ with $H\text{-}NS_{90-136}$ and by analysing an ensemble of model structures of $StpA_{91-134}$ and NOE restraints using the program Procheck. Procheck provides detailed information regarding the quality of the ensemble of structures, and was used to systematically identify incorrect assignments. Once identified, incorrect assignments were excluded from the next iteration of $StpA_{91-134}$ structure calculation. Gradually, as more unambiguous restraints were created, the structures steadily improved until no further improvements in $StpA_{91-134}$ structure were possible.

Ambiguities arose during the assignment process when one NOE cross-peak could correspond to two or more ^{1}H nuclei with the same chemical shifts. In many cases, ambiguity was solved by analysis of the preliminary models of $StpA_{91-134}$ and the

structure of H-NS_{90-136}, as NOE correlations are undetectable when the distance between ^{1}H nuclei is greater than 6 Å. When there was an NOE cross peak that potentially corresponded to more than one ^{1}H nucleus within 6 Å, then the ambiguity was not be solved.

Figure 3.21 shows an ensemble of 10 polypeptide backbone traces of $StpA_{91-134}$ structures in the intermediate stages of refinement, generated using the same set of distance and TALOS dihedral restraints. Figure 3.22 shows an ensemble of 16 backbone traces of $StpA_{91-134}$ structures derived from the final iteration of structural refinement, and represents models with the least energy associated with NOE restraint violations. The structures in the ensembles have been aligned with respect to residues 10 to 48, which have [^{1}H, ^{15}N] heteronuclear NOE values of less than 0.5.

3.2.7 ARIA

An alternative, automatic method for the determination of $StpA_{91-134}$ structure using the program ARIA was employed (Linge *et al.*, 2001; Nilges and O'Donoghue, 1998; Nilges *et al.*, 1997). The backbone and side-chain resonance assignments, the unassigned ^{13}C-NOESY-HSQC and ^{15}N-NOESY-HSQC spectra, and TALOS restraints were used as inputs to the program ARIA. The distinct advantage of the automated method of using ARIA was that no initial model structure was required. In contrast, the manual assignment method outlined in Section 3.2.6 required the published H-NS_{90-136} structure as a template. All processes in ARIA are fully automated, ensuring there are no subjective errors introduced in the NOE assignment.

Figure 3.23 shows the $StpA_{91-134}$ structures calculated during the iterations of ARIA: six structures of the best energy structures, calculated in each iteration, are overlaid with respect to each other. A consistent fold common to all structures of the ensemble, similar to the fold exhibited to the homologous H-NS_{90-136} structure, emerged around the fifth iteration, and continually improved in subsequent iterations. The secondary structures of the structures of the ensemble of the eighth iteration showed little variability, suggesting that the structural refinement process by ARIA has converged.

Figure 3.21
An ensemble of 10 $StpA_{91-134}$ structures in the process of refinement generated using the program CNS, represented by their Cα traces. Residues 1 to 4 are not shown. The structures are aligned with respect to each other, using residues 10 to 48. The figure was generated using PyMOL (Delano, 2004).

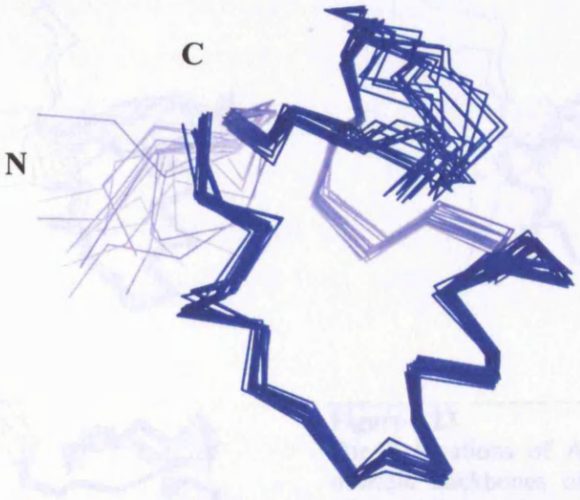


Figure 3.22
An ensemble of 16 $StpA_{91-134}$ structures generated using the program CNS, represented by tracing of the Cα nuclei. This ensemble represents the best $StpA_{91-134}$ structures that could be calculated using the manual method of NOE assignment. Residues 1 to 4 are not shown. The structures are aligned with respect to each other, using residues 10 to 48. The figure was generated using PyMOL (Delano, 2004).

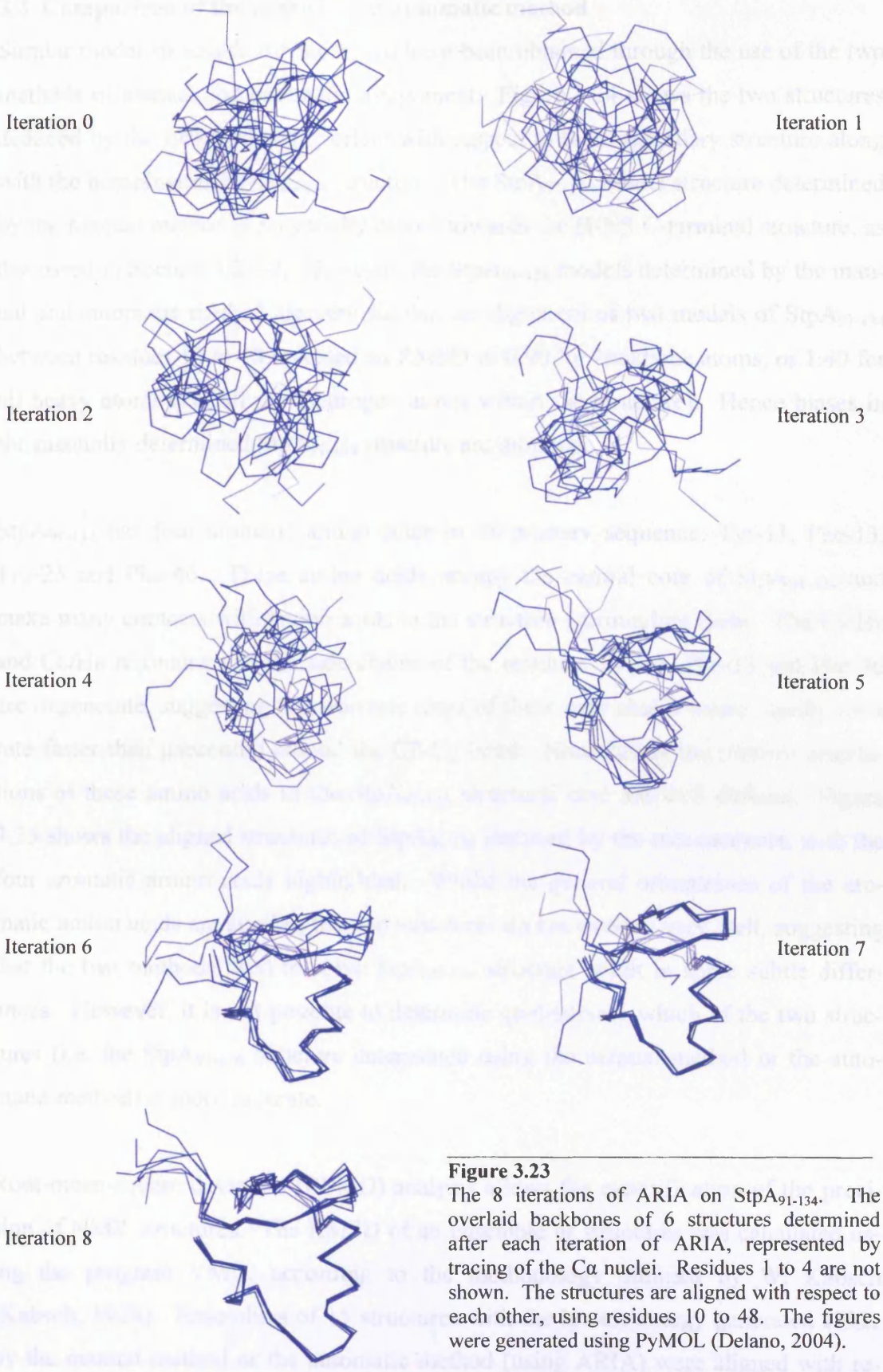


Figure 3.23
The 8 iterations of ARIA on $StpA_{91-134}$. The overlaid backbones of 6 structures determined after each iteration of ARIA, represented by tracing of the Cα nuclei. Residues 1 to 4 are not shown. The structures are aligned with respect to each other, using residues 10 to 48. The figures were generated using PyMOL (Delano, 2004).

3.3 Comparison of the manual and automatic method

Similar model structures for $StpA_{91-134}$ have been obtained through the use of the two methods of manual and automatic assignment. Figure 3.24 shows the two structures deduced by the two methods overlaid with respect to their secondary structure along with the homologous $H\text{-}NS_{90-136}$ structure. The $StpA_{91-134}$ model structure determined by the manual method is potentially biased towards the H-NS C-terminal structure, as discussed in Section 3.2.6.4. However, the $StpA_{91-134}$ models determined by the manual and automatic method are very similar: an alignment of two models of $StpA_{91-134}$ between residues 10 to 48 revealed an RMSD of 0.90 for backbone atoms, or 1.40 for all heavy atoms (i.e. all non-hydrogen atoms within the structure). Hence biases in the manually determined $StpA_{91-134}$ structure are minimal.

$StpA_{91-134}$ has four aromatic amino acids in its primary sequence: Tyr-11, Phe-13, Trp-23 and Phe-46. These amino acids occupy the central core of $StpA_{91-134}$ and make many contacts with amino acids in the structure surrounding them. The Cγ/Hγ and Cε/Hε resonances of the side chains of the residues Tyr-11, Phe-13 and Phe-46 are degenerate, suggesting the aromatic rings of these side chains rotate rapidly (at a rate faster than μseconds) around the Cβ-Cγ bond. Nonetheless the relative orientations of these amino acids in the $StpA_{91-134}$ structural core are well defined. Figure 3.25 shows the aligned structures of $StpA_{91-134}$ deduced by the two methods, with the four aromatic amino acids highlighted. Whilst the general orientations of the aromatic amino acids are similar, the two structures do not overlay very well, suggesting that the two methods used to solve $StpA_{91-134}$ structure result in some subtle differences. However, it is not possible to determine qualitatively which of the two structures (i.e. the $StpA_{91-134}$ structure determined using the manual method or the automatic method) is more accurate.

Root-mean-square deviation (RMSD) analysis allows the quantification of the precision of NMR structures. The RMSD of an ensemble of structures was calculated using the program VMD, according to the methodology outlined by W. Kabsch (Kabsch, 1978). Ensembles of 15 structures with the lowest energy generated either by the manual method or the automatic method (using ARIA) were aligned with respect to each other using residues 10 to 48. Residues with a [^{1}H, ^{15}N] heteronuclear NOE less than 0.5 are highly flexible, and consequently may readily violate sterically

forbidden conformations than the more rigid structured regions. Therefore residues 1 to 9, displaying [^{1}H, ^{15}N] heteronuclear NOEs of less than 0.5 (see Section 3.2.5), were not included in the subsequent analysis. All the structures in the ensembles were first aligned with respect to each other and the resulting ensemble averaged. Individual RMSD values were calculated by with respect to the averaged structure, and all RMSD values then averaged. RMSD values were calculated either using the backbone atoms (Cα, CO, and N) or all heavy atoms of the residues. The RMSD values are presented in Table 3.2. The RMSD values for both ensembles of $StpA_{91-134}$ structures are low, showing that they are precisely defined. The $StpA_{91-134}$ structures determined using either the manual method or automatic method could not be distinguished on the basis of the RMSD values.

Procheck was used to analyse the final sets of structures solved using the manual and automatic methods. Residues with a [^{1}H, ^{15}N] heteronuclear NOE less than 0.5 were excluded in the Procheck analysis. Figures 3.26 shows the Ramachandran plots of the backbone dihedral angles of residues 10 to 48 of 10 $StpA_{91-134}$ structures solved using the automatic method and the manual method. Due to steric hindrance, backbone dihedral angles have a limited range of φ and ψ angles, pictorially represented as 'allowable' regions on Ramachandran plots. The quality of $StpA_{91-134}$ structures can be determined by the proportion of backbone dihedral angles that lie within the allowable regions of a Ramachandran plot, as shown in Table 3.2.

Figure 3.24
A backbone alignment of the two structures of $StpA_{91-134}$ and that of $H\text{-}NS_{90-136}$. The Cα traces are shown. The backbone of the $StpA_{91-134}$ structure determined using the manual method and the automatic method are shown in light blue and dark blue, respectively. The backbone of the NMR structure of $H\text{-}NS_{90-136}$ is shown in red. The figure was generated using PyMOL (Delano, 2004).

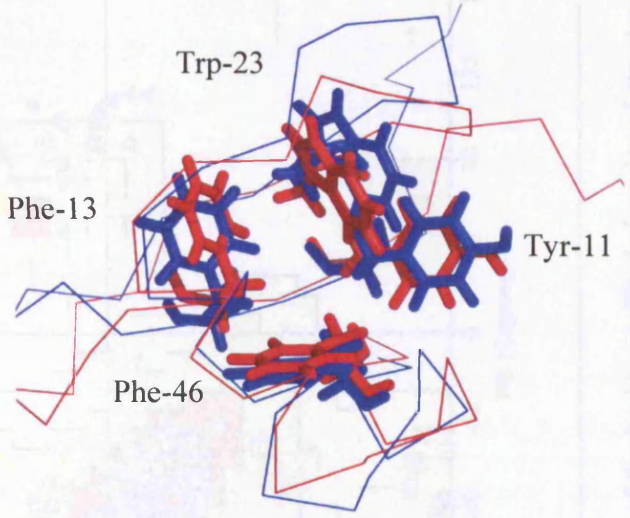


Figure 3.25
A backbone alignment of the two structures of $StpA_{91-134}$, determined by the manual method and the automatic method coloured red and blue, respectively. The residues Tyr-11, Phe-13, Trp-23 and Phe-46, which are situated in the centre of the $StpA_{91-134}$ structure, are highlighted. The figure was generated using PyMOL (Delano, 2004).

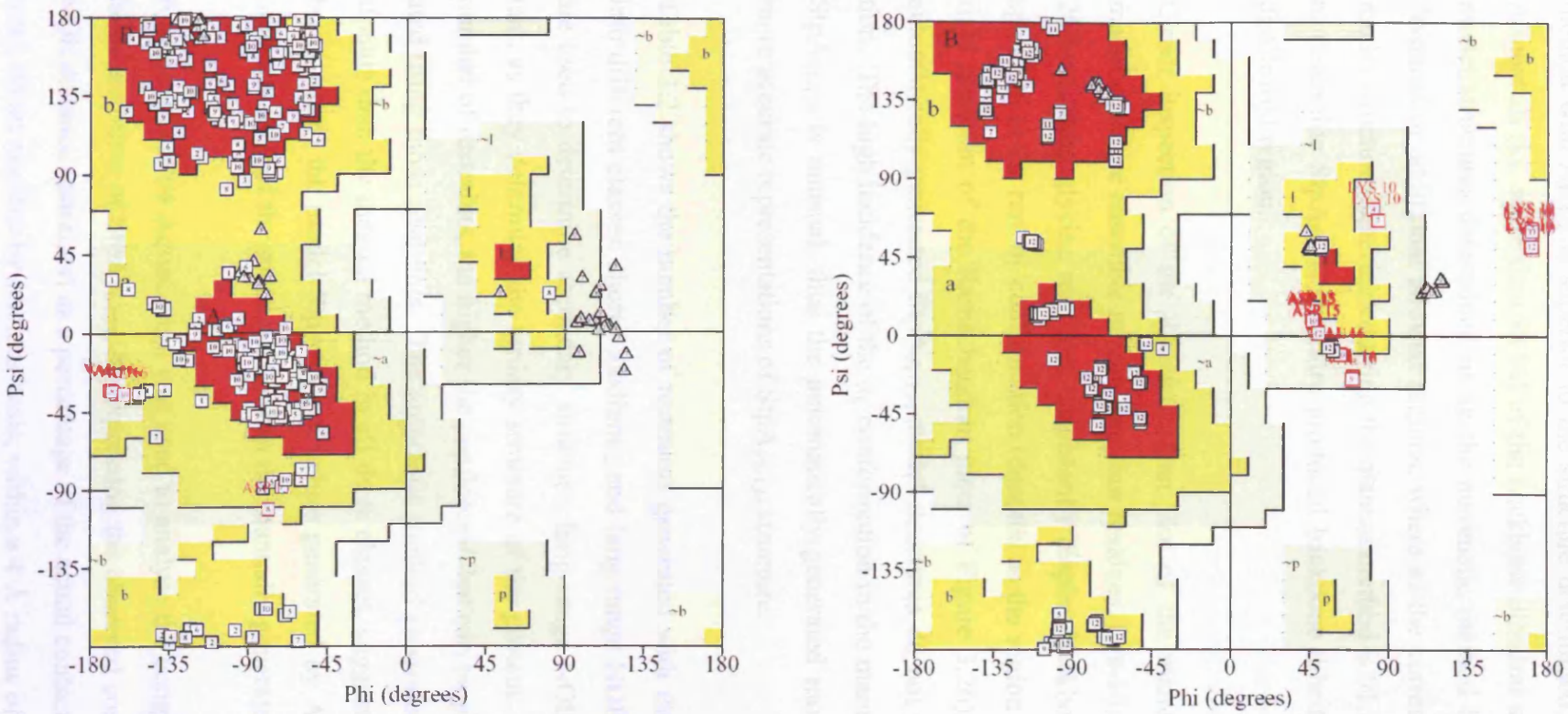


Figure 3.26

The Ramachandran plot of the backbone dihedral angles of structures of $StpA_{91-134}$ calculated using the automatic method with ARIA, generated using Procheck, is shown on the left. The corresponding Ramachandran plot of $StpA_{91-134}$ structures calculated using the manual method is shown on the right. The backbone dihedral angles of all residues are represented as squares, except glycine, which is represented as triangles. The regions shaded red represent the most sterically favoured backbone dihedral angles, the regions shaded yellow represent the additional allowed regions, and the cream-coloured regions represent the generously allowed regions. Any residues that do not lie within the most favoured (red) regions or the additional allowed (yellow) regions are highlighted in red (except glycine which exhibit different behaviour on the Ramachandran plot due to the lack of a side-chain), and constitute a violation of the limits of backbone dihedral angles imposed by steric hindrance.

These statistics suggest that the $StpA_{91-134}$ structure determined using the automatic method with ARIA is superior to the structure determined with the manual method. Almost all (i.e. more than 98 %) of the backbone dihedral angles of the 10 $StpA_{91-134}$ model structures determined using the automatic method lie within either the most favoured or additional allowed regions, where as the corresponding value for the 12 model structures determined using the manual method is 91.7%. Neither method used to determine $StpA_{91-134}$ structure produced backbone dihedral angles that lay in the disallowed region.

Closer inspection of the Ramachandran plot of the manually-generated $StpA_{91-134}$ model structure ensemble reveals that four residues, Lys-10, Asp-15, Val-16 and Gln-26 (excluding glycine residues) consistently display backbone dihedral angles corresponding to the rare α_L conformation (denoted as the region labelled 'L' in the upper-right quadrant of the Ramachandran plots of Figure 3.26). On the other hand the automatically-generated $StpA_{91-134}$ model structures do not display the α_L conformation. The high incidence of the α_L conformation in the manually-generated models of $StpA_{91-134}$ is unusual, thus the automatically-generated models are likely to be the more accurate representations of $StpA_{91-134}$ structure.

Table 3.2 shows the number of restraints generated with the two methods, separated into different classes: short-, medium-, and long-range NOEs. Medium-range NOEs are used to determine secondary structure; long-range NOEs are particularly important, as they determine the tertiary structure of the protein. Generally, the larger the number of restraints, the higher the confidence that can be placed in the model generated using those restraints. The automatic method generated more NOE distance restraints than the manual method in all three classes, suggesting (with all other things being equal) the model $StpA_{91-134}$ structure generated by ARIA is a more complete representation of the spectral data than the manually generated structure.

A program called AquaCompl was used to analyse the completeness of the NOE distance restraints of $StpA_{91-134}$, by expressing the observed contacts (in other words the NOE distance restraints) as a percentage of the actual contacts (by analysing the structure), on an nucleus-by-nucleus basis, within a 4 Å radius of that nucleus (Doreleijers *et al.*, 1999). For instance, a completeness of 100 % infers that all other nuclei within

a 4 Å shell of a nucleus *X* have a corresponding NOE distance restraint to that nucleus *X*. Figure 3.27 shows the per-residue completeness of the NOE distance restraint lists of $StpA_{91-134}$. The per-residue completeness of the structure generated by ARIA is consistently higher than the manually-generated structure. Compared with an average completeness of 58% for the NOE restraint list generated using the manual method, the average completeness of 75% for the ARIA-generated restraint list implies the data used to determine $StpA_{91-134}$ structure using ARIA is of a higher quality. Interestingly, the peaks and troughs of the per-residue completeness of both methods in Figure 3.27 match: this is a consequence of using the same NOESY-HSQC spectra to derive the restraint lists via the two methods. For example, threonine-24 and glycine-25 did not have any corresponding cross-peaks in ^{15}N-edited spectra such as [^{1}H, ^{15}N]-HSQC and ^{15}N-NOESY-HSQC: therefore the per-residue completeness for these residues is low.

Much of the analysis of the $StpA_{91-134}$ structures generated using the two methods, for example, comparison of the Ramachandran plots, the total number of restraints used, and the NOE restraint list completeness, suggests that the automatic method using ARIA produces better quality structures than the manual method. In addition, the manual method required the use of the published H-NS C-terminal domain structure to be used as a model, potentially biasing the final structure. Therefore the manually generated structure of $StpA_{91-134}$ is excluded from further analysis and discussion.

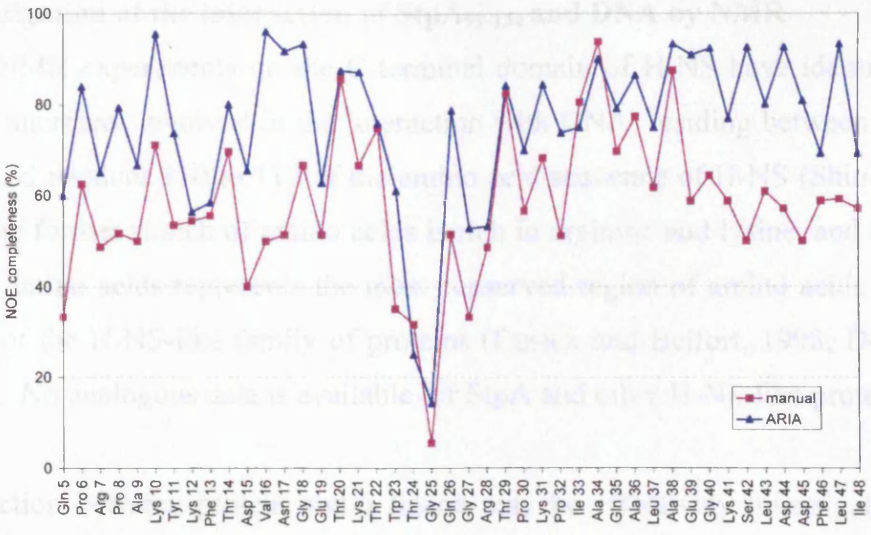


Figure 3.27
The per-residue completeness of the NOE distance restraint lists of $StpA_{91-134}$. 'Completeness' is defined as the observed contacts of a given nucleus as a percentage of the actual contacts, within a 4 Å radius of that nucleus. The restraint list generated using the manual and automatic methods are shown in red and blue, respectively. The average per-residue completeness of the restraint lists are 58% and 75%, respectively.

	Manual	Automatic
Total number of restraints	594	1177
Unambiguous	572	924
Ambiguous	-	211
Short-range	270	457
Medium-range ($1<\|i-j\|<5$)	199	324
Long-range ($\|i-j\|\geq 5$)	103	163
Dihedral (φ and ψ)	22	22
RMSD, backbone atoms (Å)	0.52 ± 0.20	0.51 ± 0.11
RMSD, all atoms (Å)	1.10 ± 0.14	1.24 ± 0.10
Average completeness of NOE completeness	58%	75%
Ramachandran plot analysis:*		
Residues in favoured region (%)	69.8 %	78.1 %
Residues in additional allowed region (%)	21.9 %	20.0 %
Residues in generously allowed regions (%)	8.3 %	1.9 %
Residues in disallowed regions (%)	0 %	0 %

Table 3.2
Structural statistics for the manual and automatic structures of $StpA_{91-134}$.
* The amino acids proline and glycine are not included in these statistics.

3.4 Investigation of the interaction of $StpA_{91-134}$ and DNA by NMR

Previous NMR experiments on the C-terminal domain of H-NS have identified two structural interfaces involved in the interaction with DNA, residing between residues 80 to 96 and residues 110 to 117 of the amino acid sequence of H-NS (Shindo *et al.*, 1999). The former stretch of amino acids is rich in arginine and lysine, and the latter stretch of amino acids represents the most conserved region of amino acids amongst members of the H-NS-like family of proteins (Cusick and Belfort, 1998; Dorman *et al.*, 1999). No analogous data is available for StpA and other H-NS-like proteins.

An interaction between protein and a ligand may be effectively investigated using NMR. A [^{1}H, ^{15}N]-HSQC spectrum displays cross-peaks that correspond to NH amide groups present in the protein sample, allowing backbone amide resonances to be examined. The chemical shift of any cross-peak in a [^{1}H, ^{15}N]-HSQC spectrum is determined by its local chemical environment, and is sensitive to any changes in pH, salt concentration, temperature and changes in structure of the protein sample (see Section 3.2.2.1). The binding of a ligand to a protein interface may lead to changes in the local chemical environment of nuclei situated on or near that protein interface. Thus the binding of a ligand to a protein can lead to qualitative changes in chemical shift of the cross-peaks corresponding to the backbone amide NH groups near or on the interaction interface. This approach was used previously to map the structural interface of the C-terminal DNA-binding domain (Shindo *et al.*, 1999).

A DNA oligonucleotide, 20 base pairs in length, of sequence CTGCACTTTAAAAA-GACGTC was titrated into a sample of [^{1}H, ^{15}N]-labelled $StpA_{91-134}$. The molar ratio of duplex DNA: $StpA_{91-134}$ was varied incrementally between 0 and 1, with [^{1}H, ^{15}N]-HSQC spectra acquired at each titration point. The overlaid spectra shown in Figure 3.28 clearly show that some cross-peaks change in chemical shift significantly during the course of the titration, whilst others do not. The total chemical shift perturbations for each individual backbone amide is calculated and plotted with respect to amino acid sequence, as shown in Figure 3.29. The majority of the change in chemical shift occurred at ratios of DNA: $StpA_{91-134}$ equal to or lower than unity: increasing the ratio above unity resulted in little or no further change in the [^{1}H, ^{15}N]-HSQC spectra (data not shown).

Many residues do not have a corresponding value in Figure 3.29. This may be due to the inherent lack of a backbone amide, for instance with Pro-6, Pro-8, Pro-30 and Pro-32, resulting in no corresponding cross-peak in the [^{1}H, ^{15}N]-HSQC. Some residues do not have a visible cross-peak corresponding to their backbone amide, for example Thr-24 and Thr-25. This may reflect high flexibility of the structural region surrounding these residues, or indicate exchange with solvent (see Section 3.2.2.1). In some cases, for instance with Arg-28 and Thr-29, the corresponding backbone amide cross-peaks lie in crowded regions of the [^{1}H, ^{15}N]-HSQC spectrum leading to badly resolved cross-peaks, and thus the chemical shift perturbations could not be calculated.

The chemical shift perturbations are classed according to magnitude: small, medium and large, with the chemical shift perturbations ranging from 0~0.1, 0.1~0.2 and 0.2~0.5 respectively. These classifications are then displayed on the model solution structure of $StpA_{91-134}$ as shown in Figure 3.30.

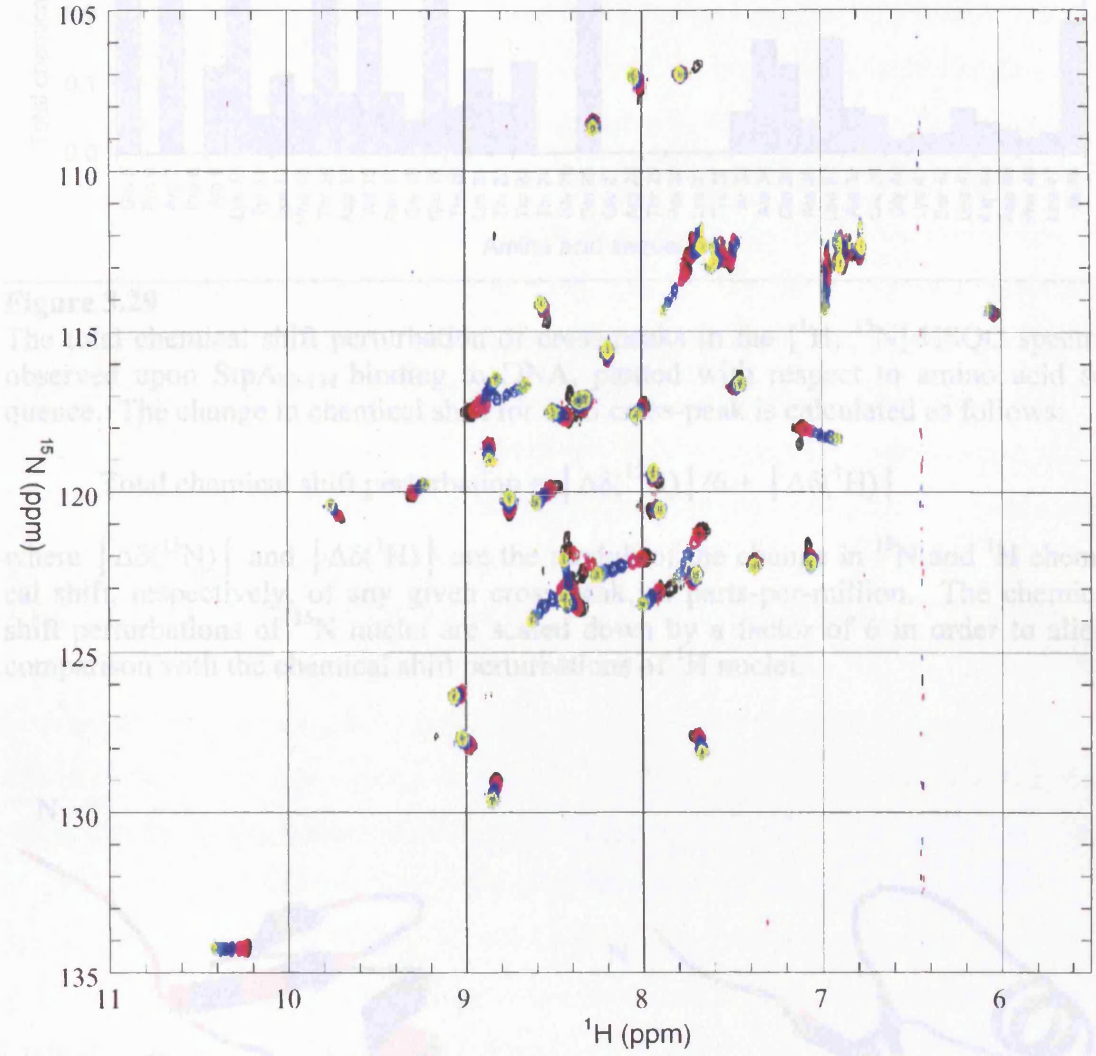


Figure 3.28
Overlaid [^{1}H, ^{15}N]-HSQC spectra of samples containing [^{1}H, ^{15}N]-labelled $StpA_{91-134}$ and various concentrations of DNA. The following molar ratios of DNA: $StpA_{91-134}$ are shown: no DNA (●), 1:10 (●), 1:5 (●), 2:5 (●), 1:2 (●), 3:5 (●), 4:5 (●), 9:10 (●), and 1:1 (●). The spectra were acquired on a 500 MHz Varian NMR spectrometer at 25 °C, using a 100 μM sample of $StpA_{91-134}$ containing 10 mM sodium phosphate pH 7.0 and 100 mM NaCl. The concentration of DNA varied between 10 μM and 100 μM.

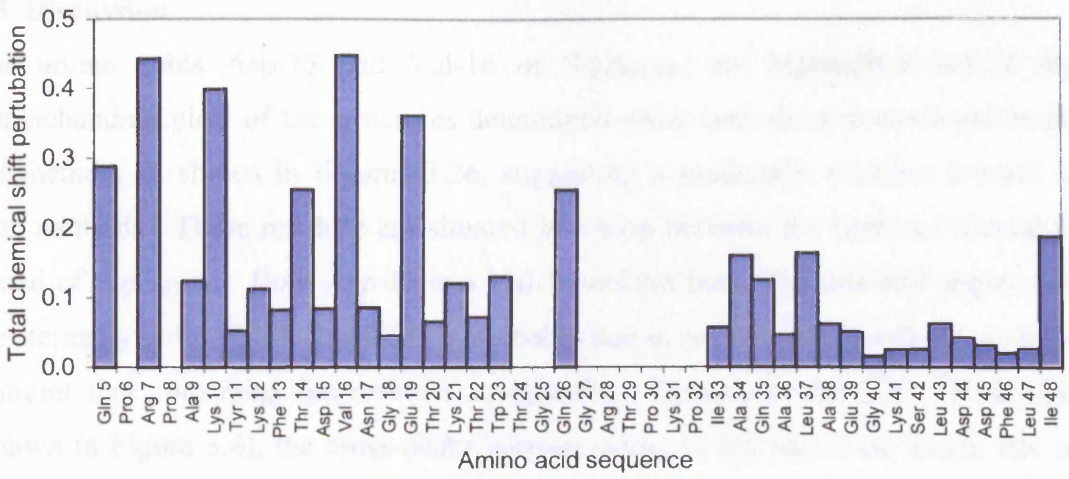


Figure 3.29
The total chemical shift perturbation of cross-peaks in the [^{1}H, ^{15}N]-HSQC spectra, observed upon StpA$_{91\text{-}134}$ binding to DNA, plotted with respect to amino acid sequence. The change in chemical shift for each cross-peak is calculated as follows:

$$\text{Total chemical shift perturbation} = |\Delta\delta(^{15}\text{N})|/6 + |\Delta\delta(^{1}\text{H})|$$

where $|\Delta\delta(^{15}\text{N})|$ and $|\Delta\delta(^{1}\text{H})|$ are the moduli of the change in ^{15}N and ^{1}H chemical shift, respectively, of any given cross-peak, in parts-per-million. The chemical shift perturbations of ^{15}N nuclei are scaled down by a factor of 6 in order to allow comparison with the chemical shift perturbations of ^{1}H nuclei.

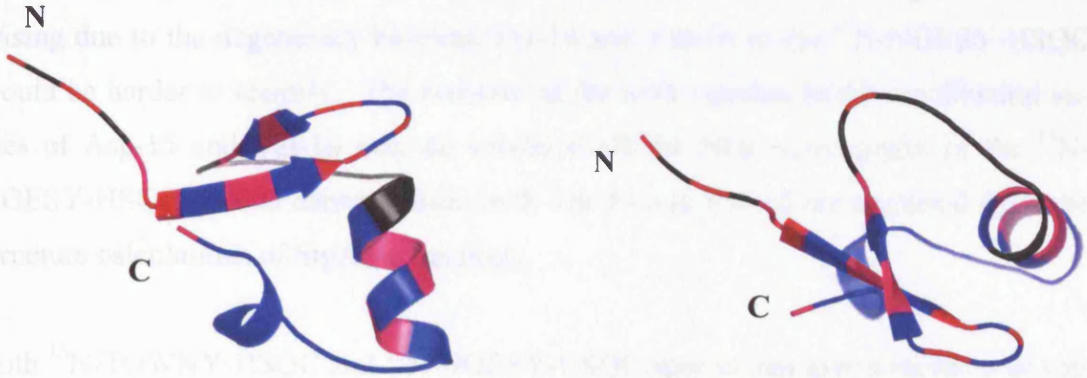


Figure 3.30
Two orthogonal representations of the model structure of StpA$_{91\text{-}134}$ solved by NMR. Residues of StpA$_{91\text{-}134}$ that experience small, medium or large chemical shift perturbations in a [^{1}H, ^{15}N]-HSQC spectrum as a result of DNA-binding are coloured blue, pink and red respectively. Residues for which there is no available data are coloured black. These figures were generated using PYMOL (Delano, 2004).

3.5 Discussion

The amino acids Asp-15 and Val-16 of $StpA_{91-134}$ are highlighted red in the Ramachandran plots of the structures determined using both the automatic and manual method, as shown in Figures 3.26, suggesting a systematic problem present in both methods. These residues are situated in a loop between the first and second β-strand of $StpA_{91-134}$. Both Asp-15 and Val-16 exhibit backbone dihedral angles that are sterically unfavourable, and this is probably due to an incorrectly defined distance restraint that constrains the residues unnaturally. Indeed, in the [^{1}H, ^{15}N]-HSQC (shown in Figure 3.4), the cross-peaks corresponding to the backbone amide HN of Thr-14 and Val-16 both overlap: therefore any spectra with pulse sequences that contain heteronuclear couplings with ^{15}N (such as [^{1}H, ^{15}N]-HSQC, CBCACONH, CCONH, ^{15}N-TOWNY-HSQC, ^{15}N-NOESY-HSQC, HNCACB, HCCONH, HNHA and HNHB) would display degeneracy between the cross-peaks corresponding to Thr-14 and Val-16. This is not particularly important for the backbone and side-chain resonance assignments, as these resonances were assigned by comparing the ^{15}N-edited spectra with other spectra (such as [^{1}H, ^{13}C]-CT-HSQC and HCCH-TOCSY), hence the chance of misassignment is small. However, as the ^{13}C-NOESY-HSQC spectrum has a lower sensitivity compared to the ^{15}N-NOESY-HSQC spectrum, comparison between the two spectra is difficult: consequently, NOE assignment errors arising due to the degeneracy between Thr-14 and Val-16 in the ^{15}N-NOESY-HSQC would be harder to identify. The problem of the unfavourable backbone dihedral angles of Asp-15 and Val-16 may be solved if all the NOE cross-peaks in the ^{15}N-NOESY-HSQC spectra deriving from both Thr-14 and Val-16 are excluded from the structure calculations of $StpA_{91-134}$ entirely.

Both ^{15}N-TOWNY-HSQC and ^{15}N-NOESY-HSQC spectra can give a measure of solvent accessibility of the backbone amide HN group: during the experiments, magnetisation generated by the pulse sequences can be directly transferred to the water by the NOE or through rapid exchange of the amide ^{1}H. Sometimes as well as seeing NOE cross-peaks to other ^{1}H nuclei in $StpA_{91-134}$ surrounding the amide HN group, solvent accessible groups show a strong NOE cross-peak to water at around 4.79 parts per million (ppm). Figure 3.31 shows the NOE cross-peaks corresponding to water at 4.79 ppm in the ^{15}N-TOWNY-HSQC spectrum (contours coloured red) and the ^{15}N-NOESY-HSQC spectrum (contours coloured blue). The residues with the amide HN

exposed to water are as follows: Glu-5, Arg-7, Ala-9, Tyr-11, Thr-14, Asp-15, Val-16, Asn-17, Gly-18, Thr-20, Lys-21, Gln-26, Glu-27, Arg-28, Thr-29, Lys-31, Gly-40, Ser-42, Leu-43 and Phe-46. Figure 3.32 shows a cartoon representation of the backbone of $StpA_{91-134}$, with the backbone amide HN groups that are exposed to water coloured red. Unsurprisingly many of the amide HN groups are exposed to solvent, as $StpA_{91-134}$ is such a small structure. The secondary structure elements, especially the helix from Pro-32 to Glu-39, provides some degree of protection to the amide HN groups from water, as they are likely to be involved in hydrogen bonding to stabilise secondary structure.

Some residues of $StpA_{91-134}$ displayed asymmetric peaks in the [^{1}H, ^{15}N]-HSQC (see Figure 3.4), suggesting these residues were involved in alternate conformations of $StpA_{91-134}$. Residues such as Lys-10, Thr-22, Lys-31, Gly-40 and Asp-45 displayed skewed peak-shapes in the [^{1}H, ^{15}N]-HSQC, suggesting that the alternate conformation is in slow exchange with the predominant conformation. Other residues such as Glu-19, Gln-26, Gly-27, Lys-31, Ala-34, Ala-36, Glu-39, Lys-41, Ser-42 and Leu-43 displayed a small, discreet peak in close proximity to the main cross-peak corresponding to the backbone amide HN, again suggesting the alternate conformation is in slow exchange. Figure 3.33 shows these residues highlighted red on $StpA_{91-134}$ structure. In some cases, the cross-peaks corresponding to the minor conformation could not be deconvolved from the larger cross-peaks corresponding to the major conformation in low resolution spectra such ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC; in other cases, although the peaks corresponding to the major and minor conformations could be deconvolved, there was no difference in the patterns of NOE between the two conformations, suggesting the differences between the major and minor conformations involve small rigid body motions of structural elements. No distinct NMR data concerning the minor conformations could be obtained, and $StpA_{91-134}$ has been modelled with only one conformation in solution in the structure calculations. Due to the small size of the StpA C-terminal domain, the minor conformations may be a consequence of the lower rigidity expected of a small polypeptide structure, rather than of some functional significance. Alternatively, as StpA binds to DNA and RNA irrespective of sequence (Azam and Ishihama, 1999; Sonnenfield *et al.*, 2001; Zhang *et al.*, 1995), the inherent flexibility may help the StpA C-terminal to bind a wide variety of DNA and RNA ligands.

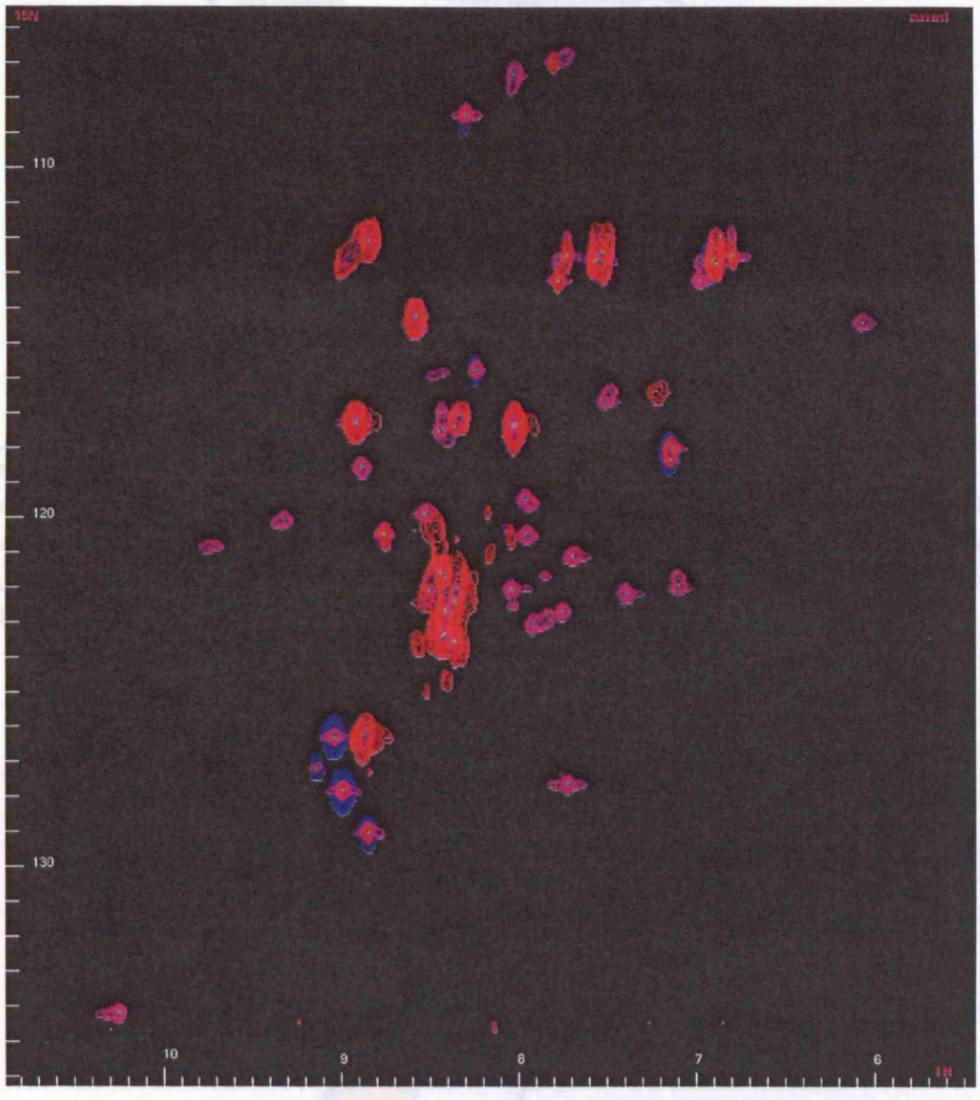


Figure 3.31
The 2D orthogonal cross-section of ^{15}N-TOWNY-HSQC and ^{15}N-NOESY-HSQC spectra of $StpA_{91-134}$, coloured red and blue, respectively, set at 4.79 ppm of the indirect ^{1}H dimension. The [^{1}H, ^{15}N]-HSQC spectrum (contours coloured pink) is also displayed for reference purposes. A strong peak at 4.79 ppm in the indirect ^{1}H dimension in either the ^{15}N-TOWNY or the ^{15}N-NOESY spectrum corresponds to water, and suggests the amide NH group is exposed to water.

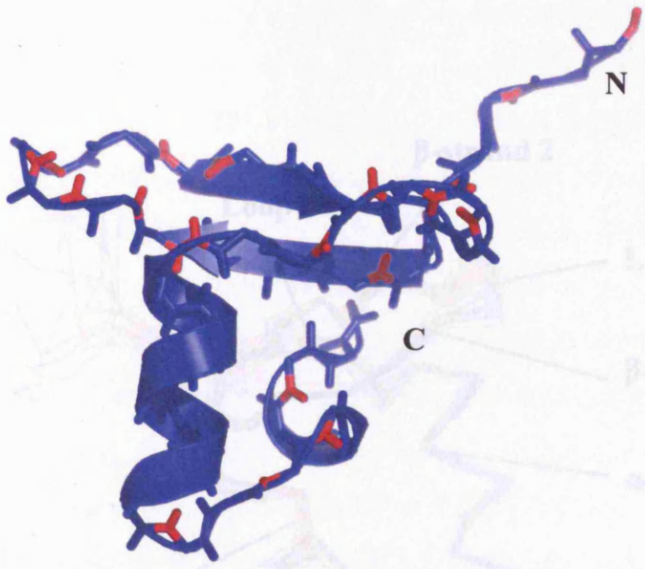


Figure 3.32
A cartoon representation of the backbone of $StpA_{91-134}$, with the exposed backbone amide NH groups highlighted in red. The $StpA_{91-134}$ structure shown was determined using the automatic method (ARIA). The figure was generated using PyMOL (Delano, 2004).

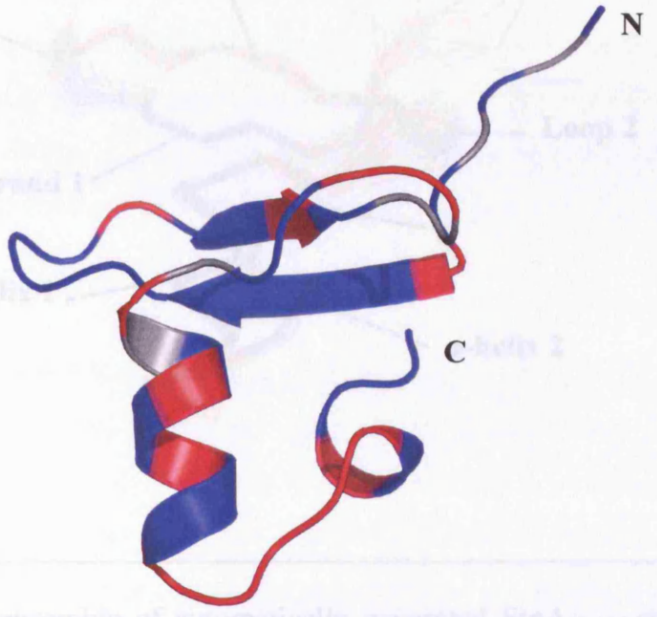


Figure 3.33
A cartoon displaying the backbone structure of $StpA_{91-134}$. Residues which display a single, symmetrical cross-peak in the [^{1}H, ^{15}N]-HSQC spectrum are coloured blue, whilst residues that display either an asymmetric cross-peak or multiple cross-peaks in the [^{1}H, ^{15}N]-HSQC, suggesting exchange between alternate conformations, are coloured red. Residues that have no corresponding peak in the [^{1}H, ^{15}N]-HSQC are shaded grey. The figure was generated using PyMOL (Delano, 2004).

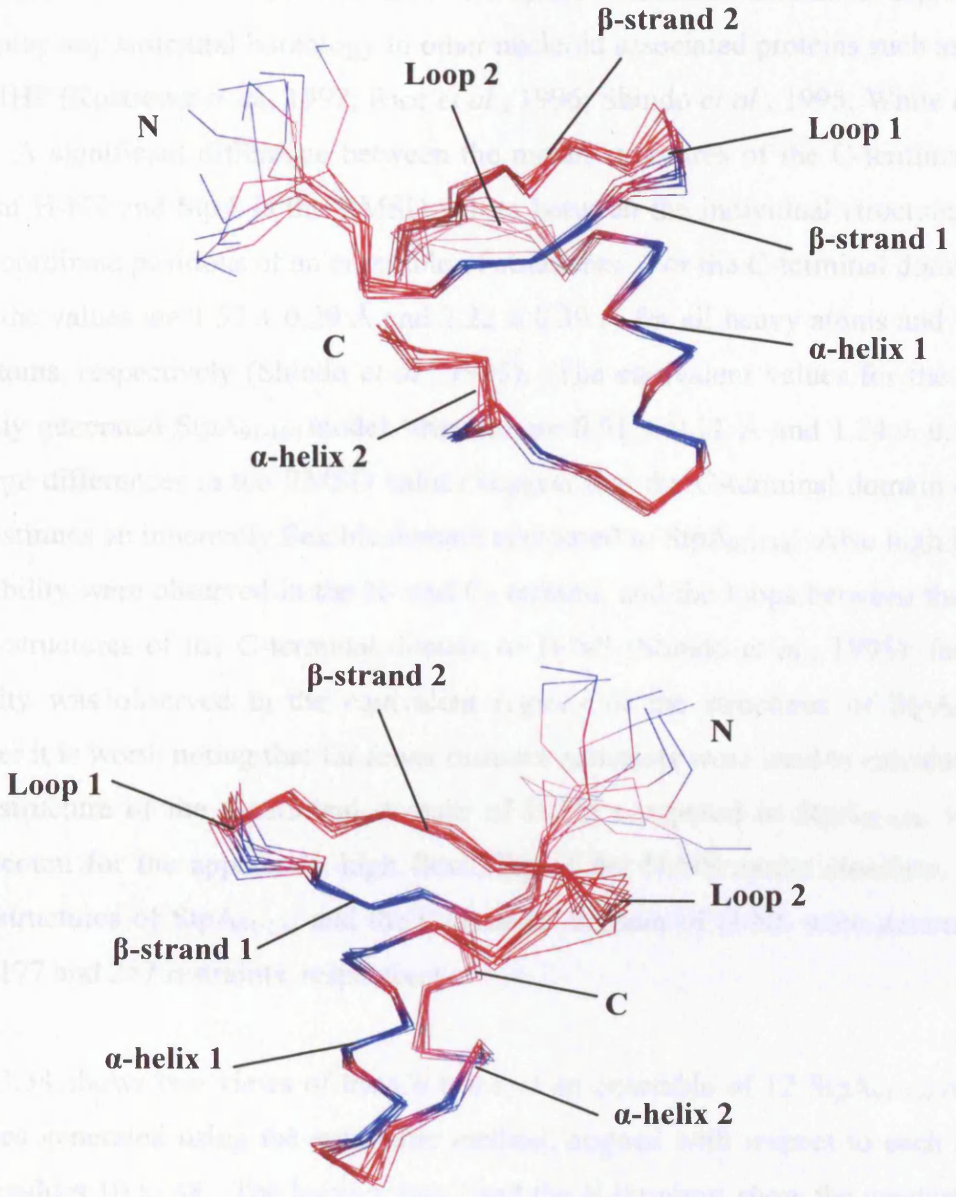


Figure 3.34
Two views of an ensemble of automatically generated $StpA_{91-134}$ structures aligned with respect to residues 10 to 48. Cα traces are shown. Residues 1 to 4 are not shown. The conserved and non-conserved residues between H-NS and StpA are coloured red and blue, respectively. The figure was generated using PyMOL (Delano, 2004)

The C-terminal domains of H-NS and StpA display similar secondary structure and tertiary fold (as shown in Figure 3.15). With 68% sequence identity (80% sequence similarity), this is not unexpected. Like H-NS, the C-terminal domain of StpA does not display any structural homology to other nucleoid associated proteins such as HU, FIS or IHF (Kostrewa *et al.*, 1992; Rice *et al.*, 1996; Shindo *et al.*, 1995; White *et al.*, 1999). A significant difference between the model structures of the C-terminal domains of H-NS and StpA is the RMSD values between the individual structures and mean coordinate positions of an ensemble of structures. For the C-terminal domain of H-NS, the values are 1.52 ± 0.29 Å and 2.22 ± 0.39 Å for all heavy atoms and backbone atoms, respectively (Shindo *et al.*, 1995). The equivalent values for the automatically generated $StpA_{91-134}$ model structure are 0.51 ± 0.11 Å and 1.24 ± 0.10 Å. The large differences in the RMSD values suggest that the C-terminal domain of H-NS constitutes an inherently flexible domain compared to $StpA_{91-134}$. Also high levels of flexibility were observed in the N- and C- termini, and the loops between the secondary structures of the C-terminal domain of H-NS (Shindo *et al.*, 1995): far less flexibility was observed in the equivalent regions of the structures of $StpA_{91-134}$. However it is worth noting that far fewer distance restraints were used to calculate the model structure of the C-terminal domain of H-NS compared to $StpA_{91-134}$, which may account for the apparently high flexibility of the H-NS model structure. The model structures of $StpA_{91-134}$ and the C-terminal domain of H-NS were determined using 1177 and 257 restraints, respectively.

Figure 3.34 shows two views of the Cα trace of an ensemble of 12 $StpA_{91-134}$ model structures generated using the automatic method, aligned with respect to each other using residues 10 to 48. The loops 1 and 2 and the N-terminus show the greatest levels of flexibility in the structures. The β-strand 2 and loop 2 represent the most conserved regions of structure between the C-terminal domain of H-NS and StpA; the β-strand 1 and the two helices seem to be the least conserved. Indeed, the loop in the C-terminal domain structure of H-NS found between residues 110 to 117 of H-NS has been shown to be the best conserved stretch of amino acids in the H-NS-like family of proteins (Cusick and Belfort, 1998; Dorman *et al.*, 1999), and has been suggested to be involved in DNA-binding in H-NS (Shindo *et al.*, 1995; Spurio *et al.*, 1997; Tippner and Wagner, 1995). The secondary structure elements, i.e. β-strand 1 and the two helices may only be required to provide structural integrity to loop 2 to allow it to ex-

ist in the right conformation to bind DNA. This suggests that point mutations that do not alter the structure of the secondary structure will be tolerated. In support of this view, there is a previous study involving random mutagenesis of H-NS failed to identify mutants with altered function as a result of mutations in the secondary structures (Ueguchi *et al.*, 1996).

A DNA oligonucleotide of 20 complimentary base pairs was titrated into a sample of $StpA_{91-134}$, and the interaction followed by multiple acquisition of [^{1}H, ^{15}N]-HSQC spectra. Upon binding DNA, the chemical shifts of many cross-peaks of the [^{1}H, ^{15}N]-HSQC spectrum shifted. These perturbations in chemical shift were quantified, and mapped onto the structure of $StpA_{91-134}$. Unfortunately, there is no apparent pattern in the chemical shift perturbations with respect to structure. Large perturbations in chemical shift, which may indicate that the corresponding backbone amide HN group is on or near the DNA-binding interface, do not apparently localise on any particular part of the $StpA_{91-134}$ structure. However there is some indication that the residues involved in and near the β-sheet structure of $StpA_{91-134}$ display the greatest chemical shift perturbations. The majority of the corresponding loop structure in $StpA_{91-134}$ found between the second β-strand and first α-helix structure (between residues Thr-24 to Thr-29) is invisible in the [^{1}H, ^{15}N]-HSQC spectra of the titration of $StpA_{91-134}$ with DNA. This is most likely to be caused by either the high level of flexibility displayed by that loop region, or rapid exchange of the ^{1}H nuclei of the backbone amide NH groups with the solvent, or both.

3.6 Backbone and side chain resonance assignment of $StpA_{91-134}$

	Residue	Nucleus	Chemical shift (ppm)		Residue	Nucleus	Chemical shift (ppm)
2	**Serine**	Hα	4.48	7	**Arginine**	HN	8.41
		Hβ#	3.82			Hα	4.58
		Cα	58.19			Hβ1	1.69
		Cβ	63.84			Hβ2	1.86
						Hγ#	1.69
3	**Histidine**	Hα	4.68			Hδ#	3.16
		Hβ1	3.12			Cα	53.57
		Hβ2	3.20			Cβ	30.28
		Hδ2	7.14			Cγ	26.85
		Hε1	8.15			Cδ	43.23
		Cα	56.08			N	122.68
		Cβ	29.92				
		Cδ2	119.93	8	**Proline**	Hα	4.39
		Cε1	137.58			Hβ1	1.83
						Hβ2	2.24
4	**Methionine**	HN	8.34			Hγ#	1.96
		Hα	4.44			Hδ1	3.58
		Hβ1	1.95			Hδ2	3.75
		Hβ2	2.02			Cα	62.38
		Hγ1	2.46			Cβ	31.86
		Hγ2	2.51			Cγ	27.23
		Cα	55.19			Cδ	50.42
		Cβ	32.81				
		Cγ	31.87	9	**Alanine**	HN	8.36
		N	122.05			Hα	3.91
						Hβ#	1.31
5	**Glutamine**	HN	8.42			Cα	52.92
		Hα	4.60			Cβ	18.98
		Hβ1	1.94			N	123.58
		Hβ2	2.07				
		Hγ#	2.40	10	**Lysine**	HN	7.69
		Hε21	6.90			Hα	4.30
		Hε22	7.55			Hβ1	1.21
		Cα	53.62			Hβ2	1.36
		Cβ	28.85			Hγ1	0.66
		Cγ	33.44			Hγ2	1.05
		N	123.41			Hδ#	1.44
		Nε2	112.61			Hε#	2.80
						Cα	56.41
6	**Proline**	Hα	4.42			Cβ	36.29
		Hβ1	1.86			Cγ	24.98
		Hβ2	2.27			Cδ	29.27
		Hγ#	2.00			Cε	41.88
		Hδ1	3.66			N	121.10
		Hδ2	3.80				
		Cα	62.79				
		Cβ	31.98				
		Cγ	27.37				
		Cδ	50.56				

	Residue	Nucleus	Chemical shift (ppm)		Residue	Nucleus	Chemical shift (ppm)
11	**Tyrosine**	HN	8.35	**15**	**Aspartate**	HN	8.99
		Hα	5.65			Hα	4.92
		Hβ1	3.01			Hβ1	2.87
		Hβ2	3.35			Hβ2	2.98
		Hδ#	6.81			Cα	52.38
		Hε#	6.52			Cβ	42.21
		Cα	56.25			N	127.79
		Cβ	42.49				
		Cδ#	133.24	**16**	**Valine**	HN	8.92
		Cε#	138.05			Hα	3.98
		N	117.40			Hβ	2.25
						Hγ1#	1.00
12	**Lysine**	HN	9.72			Hγ2#	1.03
		Hα	5.66			Cα	64.17
		Hβ#	1.76			Cβ	31.70
		Hγ#	1.40			Cγ1	20.02
		Hδ#	1.63			Cγ2	20.88
		Hε#	2.93			N	117.52
		Cα	54.53				
		Cβ	35.97	**17**	**Asparagine**	HN	8.01
		Cγ	23.93			Hα	4.78
		Cδ	29.53			Hβ1	2.91
		Cε	42.00			Hβ2	3.10
		N	120.91			Hδ21	6.98
						Hδ22	7.77
13	**Phenylanaline**	HN	8.88			Cα	52.70
		Hα	5.21			Cβ	38.32
		Hβ#	3.07			N	117.38
		Hδ#	6.80			Nδ2	113.17
		Hε#	7.08				
		Hζ	7.26	**18**	**Glycine**	HN	8.29
		Cα	55.84			Hα1	3.54
		Cβ	40.59			Hα2	4.25
		Cδ#	132.17			Cα	45.15
		Cε#	130.21			N	108.50
		Cζ	130.78				
		N	118.56	**19**	**Glutamate**	HN	8.04
						Hα	4.35
14	**Threonine**	HN	8.91			Hβ#	1.98
		Hα	4.69			Hγ1	2.12
		Hβ	3.99			Hγ2	2.23
		Hγ2#	1.22			Cα	55.73
		Cα	61.70			Cβ	30.88
		Cβ	69.32			Cγ	36.41
		Cγ2	21.75			N	122.06
		N	117.28				
				20	**Threonine**	HN	8.75
						Hα	4.57
						Hβ	3.98
						Hγ2#	1.12
						Cα	63.00
						Cβ	68.77
						Cγ2	21.89
						N	120.49

	Residue	Nucleus	Chemical shift (ppm)		Residue	Nucleus	Chemical shift (ppm)
21	**Lysine**	HN	8.85	**24**	**Threonine**	HN	8.46
		Hα	4.59			Hα	4.28
		Hβ1	0.50			Hβ	4.28
		Hβ2	1.50			Hγ2#	1.30
		Hγ1	1.16			Cα	62.65
		Hγ2	1.26			Cβ	70.53
		Hδ1	1.62			Cγ2	21.72
		Hδ2	1.76			N	120.39
		Hε#	2.96				
		Cα	53.49	**25**	**Glycine**	Hα1	0.8603
		Cβ	36.44			Hα2	3.72
		Cδ	28.69			Cα	44.22
		Cγ	24.89				
		Cε	42.39	**26**	**Glutamine**	HN	7.14
		N	129.05			Hα	4.54
						Hβ1	1.81
22	**Threonine**	HN	8.02			Hβ2	2.06
		Hα	5.65			Hγ#	2.18
		Hβ	4.36			Hε21	6.88
		Hγ2#	1.15			Hε22	7.51
		Cα	59.58			Cα	54.32
		Cβ	72.31			Cβ	30.76
		Cγ2	21.27			Cγ	33.64
		N	107.33			N	118.08
						Nε2	112.83
23	**Tryptophan**	HN	9.32				
		Hα	5.78	**27**	**Glycine**	HN	8.84
		Hβ1	3.28			Hα1	3.82
		Hβ2	3.36			Hα2	4.27
		Hδ1	8.07			Cα	44.88
		Hε1	10.25			N	112.07
		Hε3	6.99				
		Hζ2	7.08	**28**	**Arginine**	HN	8.42
		Hζ3	6.55			Hα	4.10
		Hη2	7.09			Hβ#	1.67
		Cα	55.01			Hγ1	1.56
		Cβ	32.98			Hγ2	1.69
		Cδ1	127.69			Hδ#	3.16
		Cε3	120.39			Cα	56.20
		Cζ2	136.06			Cβ	30.37
		Cζ3	122.21			Cγ	26.94
		Cη2	126.26			Cδ	43.28
		N	120.12			N	121.58
		Nε1	134.20				
				29	**Threonine**	HN	8.35
						Hα	3.22
						Hβ	3.57
						Hγ2#	0.62
						Cα	60.44
						Cβ	68.93
						Cγ2	20.97
						N	122.43

	Residue	Nucleus	Chemical shift (ppm)		Residue	Nucleus	Chemical shift (ppm)
30	**Proline**	Hα	3.85	**34**	**Alanine**	HN	7.91
		Hβ1	0.25			Hα	3.71
		Hβ2	1.67			Hβ#	1.18
		Hγ1	0.73			Cα	55.30
		Hγ2	1.19			Cβ	17.93
		Hδ1	1.27			N	123.07
		Hδ2	1.83				
		Cα	62.06	**35**	**Glutamine**	HN	8.43
		Cβ	32.74			Hα	3.92
		Cγ	27.40			Hβ1	1.91
		Cδ	48.74			Hβ2	2.04
						Hγ1	2.29
31	**Lysine**	HN	8.87			Hγ2	2.35
		Hα	3.92			Hε21	6.79
		Hβ#	1.84			Hε22	7.72
		Hγ1	1.35			Cα	58.34
		Hγ2	1.48			Cβ	28.20
		Hδ#	1.65			Cγ	33.60
		Hε#	2.94			N	117.48
		Cα	60.85			Nε2	112.51
		Cβ	29.86				
		Cγ	25.00	**36**	**Alanine**	HN	7.11
		Cδ	29.06			Hα	4.27
		Cε	41.88			Hβ#	1.70
		N	126.32			Cα	54.56
						Cβ	18.78
32	**Proline**	Hα	4.28			N	121.88
		Hβ1	1.37				
		Hβ2	2.56	**37**	**Leu**	HN	8.53
		Hγ1	2.05			Hα	4.28
		Hγ2	2.19			Hβ1	1.55
		Hδ1	2.78			Hβ2	1.75
		Hδ2	3.80			Hγ	1.87
		Cα	65.44			Hδ#	0.74
		Cβ	31.88			Cα	57.33
		Cγ	29.24			Cβ	40.55
		Cδ	50.03			Cγ	27.08
						Cδ#	22.68
33	**Isoleucine**	HN	6.07			N	119.91
		Hα	3.02				
		Hβ	0.85	**38**	**Alanine**	HN	7.75
		Hγ11	-0.66			Hα	4.20
		Hγ12	0.31			Hβ#	1.52
		Hδ1#	0.15			Cα	54.20
		Hγ2#	0.10			Cβ	17.70
		Cα	63.81			N	122.68
		Cβ	36.79				
		Cγ1	27.19				
		Cγ2	16.35				
		Cδ1	12.70				
		N	114.39				

	Residue	Nucleus	Chemical shift (ppm)		Residue	Nucleus	Chemical shift (ppm)
39	**Glutamate**	HN	7.50	**43**	**Leucine**	HN	9.03
		Hα	4.39			Hα	4.19
		Hβ1	2.19			Hβ1	1.69
		Hβ2	2.29			Hβ2	1.76
		Hγ1	2.32			Hγ	1.74
		Hγ2	2.53			Hδ1#	0.69
		Cα	55.84			Hδ2#	0.75
		Cβ	29.77			Cα	58.00
		Cγ	36.09			Cβ	41.19
		N	116.62			Cγ	26.96
						Cδ1	24.16
40	**Glycine**	HN	7.80			Cδ2	24.25
		Hα1	3.78			N	126.30
		Hα2	4.41				
		Cα	45.36	**44**	**Aspartate**	HN	8.23
		N	106.96			Hα	4.57
						Hβ1	2.60
41	**Lysine**	HN	7.95			Hβ2	2.77
		Hα	4.47			Cα	56.31
		Hβ1	1.56			Cβ	40.00
		Hβ2	2.17			N	115.79
		Hγ1	1.51				
		Hγ2	1.60	**45**	**Aspartate**	HN	7.95
		Hδ#	1.69			Hα	4.40
		Hε#	2.94			Hβ1	2.36
		Cα	55.80			Hβ2	2.70
		Cβ	31.94			Cα	56.65
		Cγ	25.75			Cβ	40.14
		Cδ	28.72			N	119.50
		Cε	42.15				
		N	120.52	**46**	**Phenylanaline**	HN	8.32
						Hα	4.87
42	**Serine**	HN	8.57			Hβ1	3.22
		Hα	4.83			Hβ2	4.17
		Hβ1	3.81			Hδ#	7.84
		Hβ2	3.98			Hε#	7.27
		Cα	56.54			Hζ	7.12
		Cβ	65.22			Cα	58.17
		N	114.24			Cβ	39.31
						Cδ#	132.87
						Cε#	130.89
						Cζ	129.18
						N	117.14

	Residue	Nucleus	Chemical shift (ppm)
47	**Leucine**	HN	7.40
		Hα	4.42
		Hβ1	1.61
		Hβ2	1.81
		Hγ	1.93
		Hδ1#	0.77
		Hδ2#	1.00
		Cα	56.53
		Cβ	42.91
		Cγ	27.40
		Cδ1	24.26
		Cδ2	25.41
		N	122.17
48	**Isoleucine**	HN	7.72
		Hα	3.73
		Hβ	1.45
		Hγ11	0.90
		Hγ12	1.41
		Hγ2#	0.85
		Hδ1#	0.73
		Cα	64.75
		Cβ	39.19
		Cγ1	29.42
		Cγ2	17.75
		Cδ1	14.33
		N	127.62

Chapter 4

The oligomeric properties of H-NS and StpA

4.1 Introduction

In this chapter, the oligomeric properties of H-NS and StpA are investigated. In addition the temperature dependence of H-NS oligomerisation is explored. The similarities in structure and domain organisation of H-NS and StpA are established. Results that support the 'head-to-tail' model of the functional oligomerisation of H-NS and StpA are presented, generated from a variety of biophysical techniques such as circular dichroism spectroscopy (CD), nuclear magnetic resonance spectroscopy (NMR) and size-exclusion chromatography (SEC). Similarities and differences of the mechanism of higher order oligomerisation of H-NS and StpA are highlighted. Finally, the interaction between StpA and H-NS is characterised with SEC experiments.

Members of the H-NS-like family of proteins have been suggested to confer the ability to adapt to a variety of ecological challenges (Tendeng and Bertin, 2003). With a recent study characterising the three way interaction of three members of the H-NS-like family of proteins, H-NS, StpA and Sfh in *Shigella flexneri* strain 2457T, and the identification of five members of the H-NS-like family of proteins in *Pseudomonas putida*, clearly the interaction between members of the H-NS-like family is biologi-

cally significant (Deighan *et al.*, 2003; Beloin *et al.*, 2003; Tendeng and Bertin, 2003). Study of the interaction between H-NS and StpA presents an opportunity to investigate the mechanisms by which the H-NS-like family of proteins exert their regulatory effects in the cell.

4.2 Changes in temperature affect the oligomeric state of H-NS

The dependence of the oligomerisation of H-$NS_{1\text{-}89}$ on temperature was investigated by SEC using a Superdex 75 XK16/60 column (Amersham Biosciences) equipped with a thermostatic jacket. Samples containing 168 μM H-$NS_{1\text{-}89}$ were analysed at different temperatures between 17.5°C and 45°C, with protein elution detected by absorbance at 280 nm, as shown in Figure 4.1. All H-$NS_{1\text{-}89}$ samples were incubated at the relevant temperature for one hour prior to analysis with SEC.

H-$NS_{1\text{-}89}$ yields an asymmetric SEC trace, with a sharp leading edge and a broad extended tail, characteristic of self-associating molecules (Stevens, 1986; Stevens, 1989). The elution volume of the largest oligomeric species in a sample of H-$NS_{1\text{-}89}$, given by the peak maximum of the SEC trace, varies with respect to temperature. As the temperature is increased, the elution volume of the largest oligomeric species in a sample of H-$NS_{1\text{-}89}$ increases, suggesting that the sizes of the largest oligomeric species of H-$NS_{1\text{-}89}$ decreases with increasing temperature.

There is a small peak in all of the SEC traces shown in Figure 4.1 with an elution volume of 44.5 ml, which is apparently insensitive to changes in temperature. The void volume of the column is 44.4 ml (see Section 2.7.3). The exclusion limit of the Superdex 75 column is approximately 100 kDa. Therefore the small peak with an elution volume of 44.5 ml is likely to correspond to impurities or aggregated protein carried through the purification process of H-$NS_{1\text{-}89}$ with a molecular weight in excess of 100 kDa, and can be ignored in the analysis of the data.

In Figure 4.1 there is a small peak with an elution volume of 71.5 ml, whose peak-height increases with increasing temperature. The peak has the largest area at 45°C, compared to lower temperatures; at 45°C, the area of this peak is approximately 15 % of the total area under SEC trace corresponding to the total H-$NS_{1\text{-}89}$ sample. According to molecular weight standards, this peak corresponds to a molecular weight of

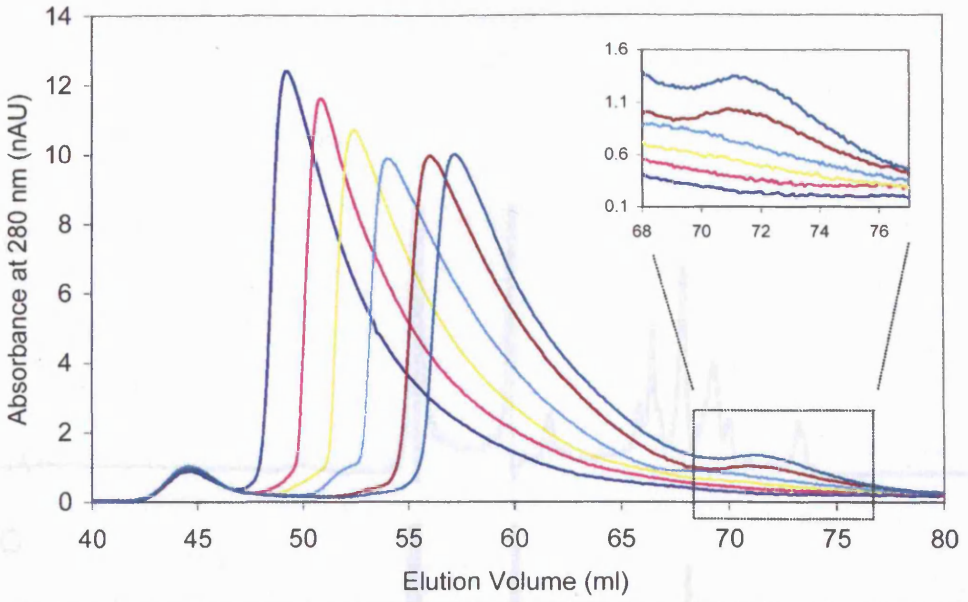


Figure 4.1
SEC traces of H-NS_{1-89} at different temperatures. 2 ml samples containing 168 μM H-NS_{1-89}, 20 mM sodium phosphate pH 7.0 and 300 mM NaCl were injected into a Superdex 75 XK16/60 column with a flow rate of 1.5 ml/min. Protein elution was detected by absorbance at 280 nm. The samples were analysed at the following temperatures: 17.5 °C (—), 25 °C (—), 30 °C (—), 35 °C (—), 40 °C (—), and 45 °C (—). The inset shows an expansion of the region of the SEC traces corresponding to elution volumes between 68 ml and 76 ml.

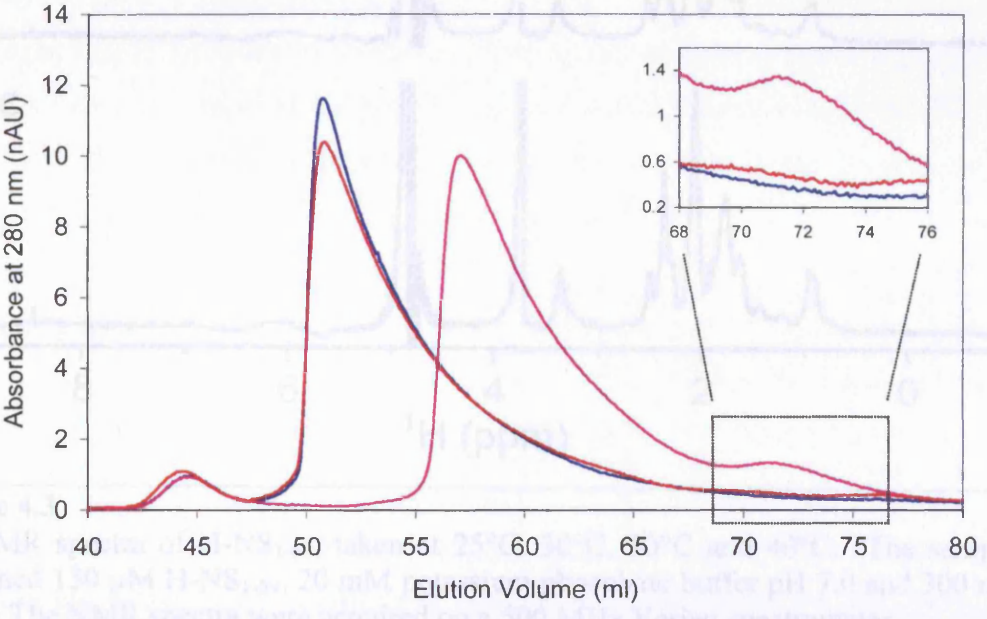


Figure 4.2
SEC traces of H-NS_{1-89}, showing the reversibility of the temperature induced changes. A control sample of H-NS_{1-89} was incubated overnight at 45°C, incubated at 4°C for 6 hours, incubated at 25°C for one hour and then analysed by SEC at 25°C (—). Also shown are the SEC traces of samples of H-NS_{1-89} ran at 25° (—) and 45°C (—).

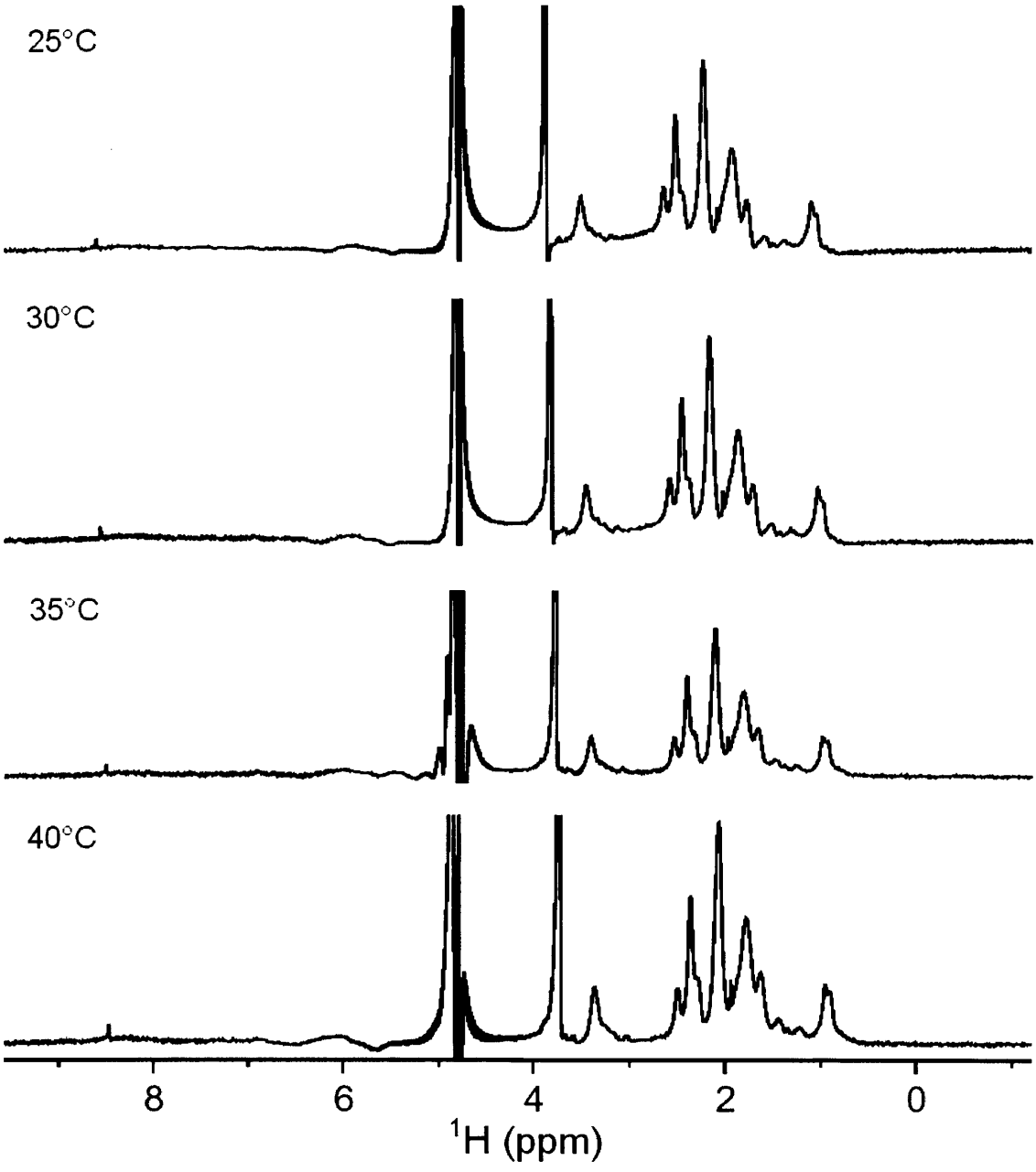


Figure 4.3
1D NMR spectra of H-NS_{1-89}, taken at 25°C, 30°C, 30°C and 40°C. The samples contained 150 μM H-NS_{1-89}, 20 mM potassium phosphate buffer pH 7.0 and 300 mM NaCl. The NMR spectra were acquired on a 500 MHz Varian spectrometer.

approximately 25 kDa, suggesting this peak may correspond to a dimer of H-NS_{1-89} (with a MW of 21 kDa). However, it is worth noting that the H-NS_{1-64} dimer does not migrate in SEC columns in a comparable fashion to globular protein standards (see Section 4.3.6). With a molecular weight of 15.9 kDa, dimeric H-NS_{1-64} gives a symmetric peak with an elution volume of 70.0 ml (data not shown). As the small peak in Figure 4.1 elutes at 71.5 ml, it may correspond to a species that is smaller than a dimer of H-NS_{1-64}, i.e. a monomer of H-NS_{1-89}.

The temperature-induced changes in H-NS_{1-89}, i.e. the shift in elution volume of the largest oligomeric species and the appearance of the small peak at 71.5ml, are reversible. A control sample of H-NS_{1-89} was incubated overnight at 45°C, incubated at 4°C for 6 hours, incubated at 25°C for one hour and then analysed by SEC at 25°C, as shown in Figure 4.2. This SEC trace was compared to a SEC trace of H-NS_{1-89} run at 25°C and 45°C. The control sample (incubated at 45°C, 4°C, 25°C and then analysed at 25°) does not display any of the characteristics of H-NS_{1-89} at 45°C, such as higher elution volume of the peak maximum, or the small peak at 71.5 ml elution volume.

4.2.1 Temperature-induced changes in H-NS_{1-89} are not detectable by NMR

1D NMR spectra of H-NS_{1-89} were acquired in order to assess the affect of temperature on the oligomeric state of H-NS_{1-89}. Several factors impose upper limits on the molecular weight of proteins that may be studied by NMR spectroscopy (see Section 3.1.3), such that proteins and complexes in excess of 30 kDa cannot be analysed by 1D NMR, as the majority of the signals will be invisible.

H-NS_{1-89} is known to oligomerise in a concentration-dependent manner, to form oligomeric complexes with molecular weights in excess of 30 kDa (Smyth *et al.*, 2000). As seen in Figure 4.3, a 1D NMR spectrum of H-NS_{1-89} at 25°C displays only a few peaks. Some of the peaks cluster around 2 ppm on the ^{1}H chemical shift scale, and are likely to correspond to methyl groups present in H-NS_{1-89}. There are little or no peaks corresponding to Hα or Hβ nuclei (usually found around 3 to 4 ppm on the ^{1}H chemical shift scale). There are no peaks corresponding to backbone amide ^{1}H nuclei (usually found around 5 to 8 ppm). There is a very small peak found at 8.5 ppm, and probably corresponds to ^{1}H nuclei resonances from an aromatic side chain in H-NS_{1-89}. The distinct lack of ^{1}H nuclei resonance peaks at 25°C is consistent with the find-

ing that H-$NS_{1\text{-}89}$ forms vast oligomeric forms in excess of 30 kDa. 1D NMR spectra were acquired at different temperatures of 30°C, 35°C and 40°C (also shown in Figure 4.3). There are no observable differences between any of the spectra acquired at different temperatures, suggesting temperature-induced changes in the oligomeric state are not detectable by NMR.

According to comparisons with globular molecular weight standards, the small peak at an elution volume of 71.5 ml that appears in the SEC traces of H-$NS_{1\text{-}89}$ at higher temperatures (see Figure 4.1) corresponds to a molecular weight below 30 kDa, suggesting detection of the species corresponding to this small peak is amenable to study by NMR. However, no extra peaks are visible at elevated temperature in the 1D NMR spectra. The concentration of the species corresponding to the small peak is below the detection capabilities of the NMR spectrometer.

4.3 The biophysical characterisation of StpA

4.3.1 StpA has a similar domain structure to H-NS

The domain structure of $StpA_{FL}$ was investigated using NMR spectroscopy. [^{1}H, ^{15}N]-HSQC spectra of $StpA_{FL}$ and $StpA_{91\text{-}134}$ were acquired, as shown in Figure 4.4. The [^{1}H, ^{15}N]-HSQC spectrum of $StpA_{91\text{-}134}$ overlays that of $StpA_{FL}$ almost perfectly. All cross-peaks observed in the [^{1}H, ^{15}N]-HSQC spectrum of $StpA_{91\text{-}134}$ are present in the corresponding [^{1}H, ^{15}N]-HSQC spectrum of $StpA_{FL}$. The chemical shifts of the cross-peaks common in both [^{1}H, ^{15}N]-HSQC spectra are almost identical.

The chemical shift of a given cross peak in a NMR spectrum is determined by its local chemical environment and is sensitive to pH, temperature, salt concentration and, most importantly, protein structure (see Section 3.1.1). The comparison of the two [^{1}H, ^{15}N]-HSQC spectra suggests that the residues 91 to 134 of $StpA_{FL}$ form an independent globular domain regardless of the presence or absence of the first N-terminal 90 residues. This suggests that $StpA_{91\text{-}134}$ forms a domain that is truly structurally independent from the rest of the $StpA_{FL}$ protein (i.e. residues 1 to 90). This observation is matched in many other studies on H-NS.

There are several factors that impose upper limits on the molecular weights of proteins that can be studied by solution nuclear magnetic resonance (NMR) spectroscopy.

These include rapid relaxation of transverse magnetisation, broadening signal line-widths and resonance overlap, all of which are consequences of studying proteins of high molecular weight by solution NMR (see Section 3.1.3). In the case of a [^{1}H, ^{15}N]-HSQC experiment, the upper limit of protein molecular weight is around 30 kDa, and signals deriving from proteins or protein complexes larger than 30 kDa generally will be invisible. However in some cases, even when the protein or protein complex is larger than 30 kDa, peaks deriving from local regions of high mobility and flexibility are clearly visible in a NMR spectrum. The higher mobility exhibited by these regions mean that the transverse magnetisation relaxes much more slowly, leading to conditions much more amenable to NMR spectroscopy. Hence in a [^{1}H, ^{15}N]-HSQC spectrum of a protein larger than 30 kDa, one might expect to observe nothing other than a few cross peaks deriving from the residues found in loops and termini.

On the basis of the high levels of sequence and structural homology between H-NS and StpA, the N-terminal domain of StpA is likely to mediate the formation of a large, polydisperse oligomeric complex (Williams *et al.*, 1996; Cusick and Belfort, 1998; Johansson and Uhlin, 1999; Smyth *et al.*, 2000; Johansson *et al.*, 2001). This complex is likely to be larger than 30 kDa; for instance, dimeric $StpA_{FL}$, the smallest native form of $StpA_{FL}$, has a molecular weight of 31 kDa. The NMR signals from such a large complex formed from the oligomerised N-terminal domains of $StpA_{FL}$ cannot be observed in a [^{1}H, ^{15}N]-HSQC spectrum. However the C-terminal domain of $StpA_{FL}$, corresponding to residues 91 to 134, is linked to the N-terminal domain by a long flexible linker (Cusick and Belfort, 1998). Consequently the C-terminal domain exhibits a higher degree of mobility in solution than the N-terminal domain, giving rise to an observable NMR spectrum: indeed the entire C-terminus is visible in a [^{1}H, ^{15}N]-HSQC spectrum of $StpA_{FL}$.

The [^{1}H, ^{15}N]-HSQC of $StpA_{FL}$ displays between 8 and 10 extra peaks, clustered around the random coil region of the spectrum (i.e. at around a chemical shift of 8.0 ppm to 8.5 ppm on the ^{1}H scale). These cross peaks could correspond either to the recombinant N-terminal His-tag or some residues in the linker region of $StpA_{FL}$. An attempt at assigning these cross peaks by identifying amide backbone connectivities with a 2D-^{15}NOESY-HSQC experiment was unsuccessful (data not shown).

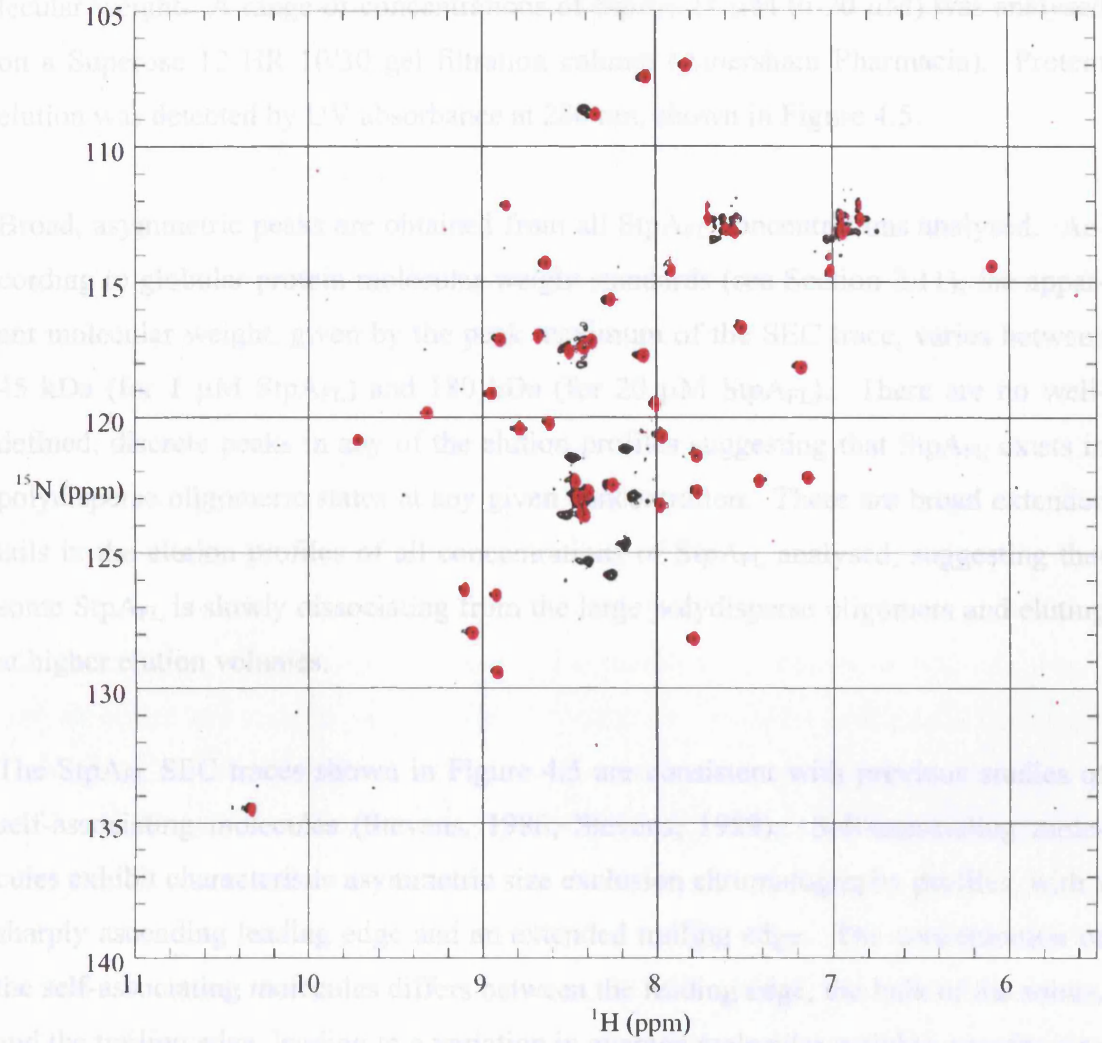


Figure 4.4
Overlaid [^{1}H, ^{15}N]-HSQC spectra of $StpA_{FL}$ and $StpA_{91\text{-}134}$, with contours coloured black and red respectively. Both the samples of $StpA_{FL}$ (at 1 mM concentration) and $StpA_{91\text{-}134}$ (at 200 µM concentration) were uniformly labelled with ^{15}Nitrogen (^{15}N), buffered in 20 mM sodium phosphate pH 7.0, 1 M NaCl, 10 % D_2O (v/v).

4.3.2 $StpA_{FL}$ oligomerises in a concentration dependent manner

Due the high sequence homology and structural similarities of StpA and H-NS (Cusick and Belfort, 1998), the oligomerisation properties of StpA were of interest. Size exclusion chromatography (SEC) experiments were conducted on $StpA_{FL}$ in order to determine the effect of differing concentrations of $StpA_{FL}$ on its apparent molecular weight. A range of concentrations of $StpA_{FL}$ (1 μM to 20 μM) was analysed on a Superose 12 HR 10/30 gel filtration column (Amersham Pharmacia). Protein elution was detected by UV absorbance at 280 nm, shown in Figure 4.5.

Broad, asymmetric peaks are obtained from all $StpA_{FL}$ concentrations analysed. According to globular protein molecular weight standards (see Section 2.11), the apparent molecular weight, given by the peak maximum of the SEC trace, varies between 45 kDa (for 1 μM $StpA_{FL}$) and 180 kDa (for 20 μM $StpA_{FL}$). There are no well-defined, discrete peaks in any of the elution profiles suggesting that $StpA_{FL}$ exists in polydisperse oligomeric states at any given concentration. There are broad extended tails in the elution profiles of all concentrations of $StpA_{FL}$ analysed, suggesting that some $StpA_{FL}$ is slowly dissociating from the large polydisperse oligomers and eluting at higher elution volumes.

The $StpA_{FL}$ SEC traces shown in Figure 4.5 are consistent with previous studies of self-associating molecules (Stevens, 1986; Stevens, 1989). Self-associating molecules exhibit characteristic asymmetric size exclusion chromatography profiles, with a sharply ascending leading edge and an extended trailing edge. The concentration of the self-associating molecules differs between the leading edge, the bulk of the solute, and the trailing edge, leading to a variation in average molecular weights over the size exclusion chromatography profile. Consequently the rate of migration varies throughout the profile (Stevens, 1986). The sharply ascending leading edge is caused by the dissociation of the largest complexes migrating ahead of the main bulk solute, whilst the extended trailing edge is caused by declining rates of reassociation of the molecules in the trailing edge as the concentration of the molecules fall (Stevens, 1989).

All $StpA_{FL}$ used in this thesis is in its tagged form, i.e. the recombinant N-terminal His-tag of $StpA_{FL}$ is not proteolytically cleaved off after purification prior to use in

experiments. Under conditions tested (between 300 mM NaCl to 500 mM NaCl, between pH 6.0 to 8.0 in either 20 mM sodium phosphate buffer or 20 mM Tris-HCl where appropriate), $StpA_{FL}$ precipitated immediately after proteolytic cleavage of the N-terminal His-tag using thrombin. When the His-tag is left uncleaved by thrombin, $StpA_{FL}$ is stable in solution under appropriate conditions. The His-tag is relatively small (17 residues long) and apparently does not prevent the formation of higher order oligomers of $StpA_{FL}$.

4.3.3 The N-terminal domain of StpA is α-helix-rich

Circular dichroism (CD) experiments conducted on H-NS suggested the N-terminal domain of H-NS was predominantly α-helical, denoted by the strong maximum at 192 nm and the strong minima at 208 and 220 nm (Smyth *et al.*, 2000; Schroder *et al.*, 2001). In a separate CD spectroscopy study, $H\text{-}NS_{FL}$ was suggested to consist predominantly of α-helix (Schroder *et al.*, 2001). These observations are consistent with NMR data of the N-terminal domain of H-NS (Renzoni *et al.*, 2001; Esposito *et al.*, 2002; Bloch *et al.*, 2003).

CD spectroscopy experiments were conducted on $StpA_{1\text{-}65}$ in order to establish secondary structure and stability in a variety of conditions. Samples of $StpA_{1\text{-}65}$ (at 30 μM concentrations) were prepared in a solution buffered with 20 mM sodium phosphate at pH 7.0, with different concentrations of NaCl: 10mM, 50mM, 100mM, 300mM and 500mM. CD spectra were recorded between 180 nm and 260 nm, as shown in Figure 4.6. The spectra give strong minima at 208 nm and 220 nm and a maximum at 192 nm (where visible), suggesting $StpA_{1\text{-}65}$ is α-helix-rich. The CD spectra suggest that the secondary structure of $StpA_{1\text{-}65}$ is independent of NaCl concentration in the range 10 mM to 500 mM at 25 °C. The small variation in CD spectra of the different samples is most likely to originate from sample preparation, rather than salt dependence of CD spectra of $StpA_{1\text{-}65}$. Note that the CD data is truncated at shorter wavelengths: data points with corresponding dynode voltages higher than 600 volts are unreliable due to low signal-to-noise ratios, and have been discarded.

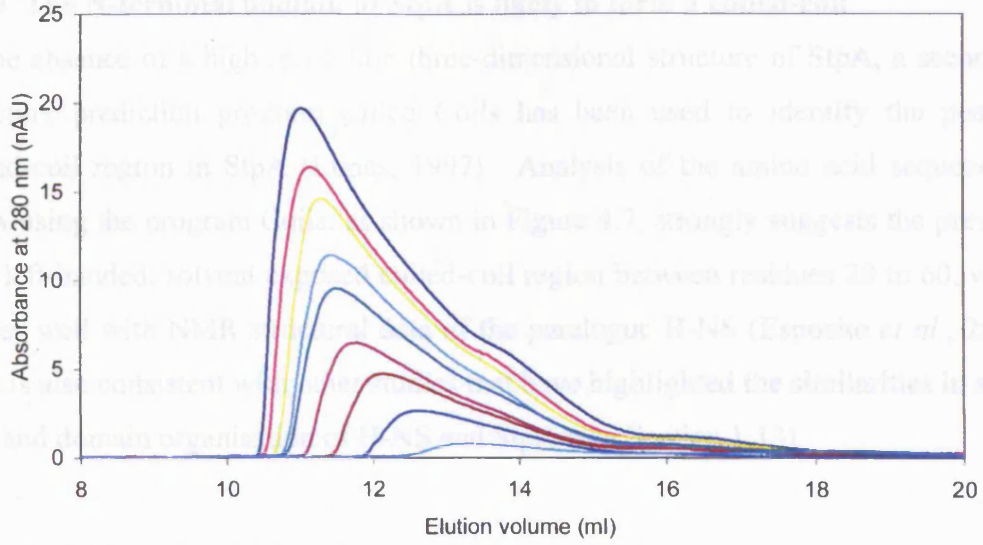


Figure 4.5
Overlaid SEC traces of $StpA_{FL}$ at different concentrations. 1 ml samples of $StpA_{FL}$ were analysed on a SEC Superose 12 10/30 column (Amersham Pharmacia) at 0.4 ml/min. The samples were buffered in 20 mM sodium phosphate, pH 7.0, 500 mM NaCl, 0.1 mM EDTA and 0.1 mM NaN_3. The following concentrations of $StpA_{FL}$ are displayed: 20 μM (—), 17 μM (—), 15 μM (—), 12 μM (—), 10 μM (—), 7 μM (—), 5 μM (—), 3 μM (—), and 1 μM (—).

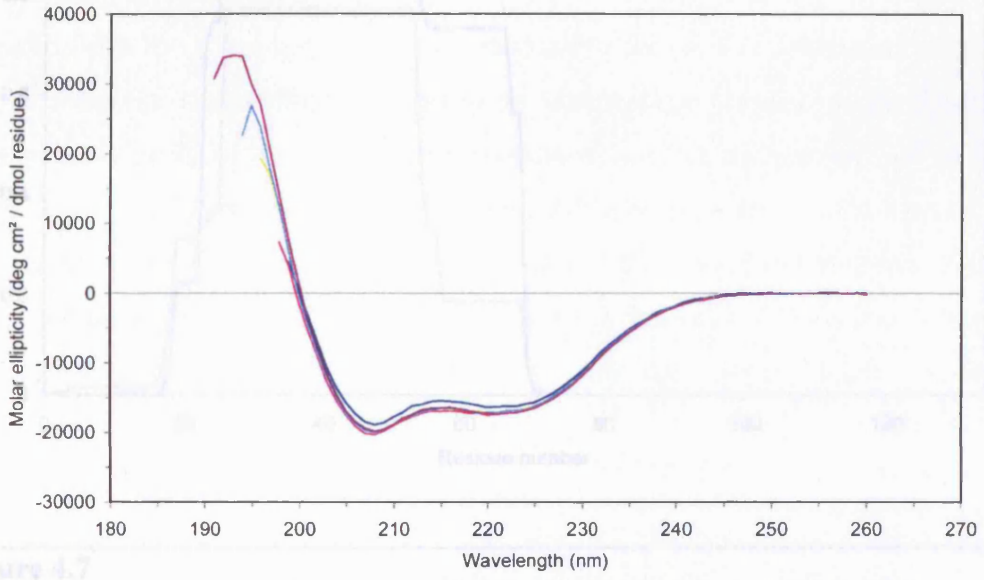


Figure 4.6
CD spectra of $StpA_{1-65}$. Samples containing 30 μM of $StpA_{1-65}$ at 500 mM NaCl (—), 300 mM NaCl (—), 100 mM NaCl (—), 50 mM NaCl (—), and 10 mM NaCl (—) were prepared in 20 mM sodium phosphate, pH 7.0. The CD spectra were recorded between 180 nm and 260 nm at 25 °C. Data points corresponding with high absorbance of the sample (when the dynode voltage was greater than 600 volts) have been discarded.

4.3.4 The N-terminal domain of StpA is likely to form a coiled-coil

In the absence of a high resolution three-dimensional structure of StpA, a secondary structure prediction program called Coils has been used to identify the possible coiled-coil region in StpA (Lupas, 1997). Analysis of the amino acid sequence of StpA using the program Coils, as shown in Figure 4.7, strongly suggests the presence of a left-handed, solvent exposed coiled-coil region between residues 20 to 60, which agrees well with NMR structural data of the paralogue H-NS (Esposito *et al.*, 2002). This is also consistent with other studies that have highlighted the similarities in structure and domain organisation of H-NS and StpA (see Section 1.13).

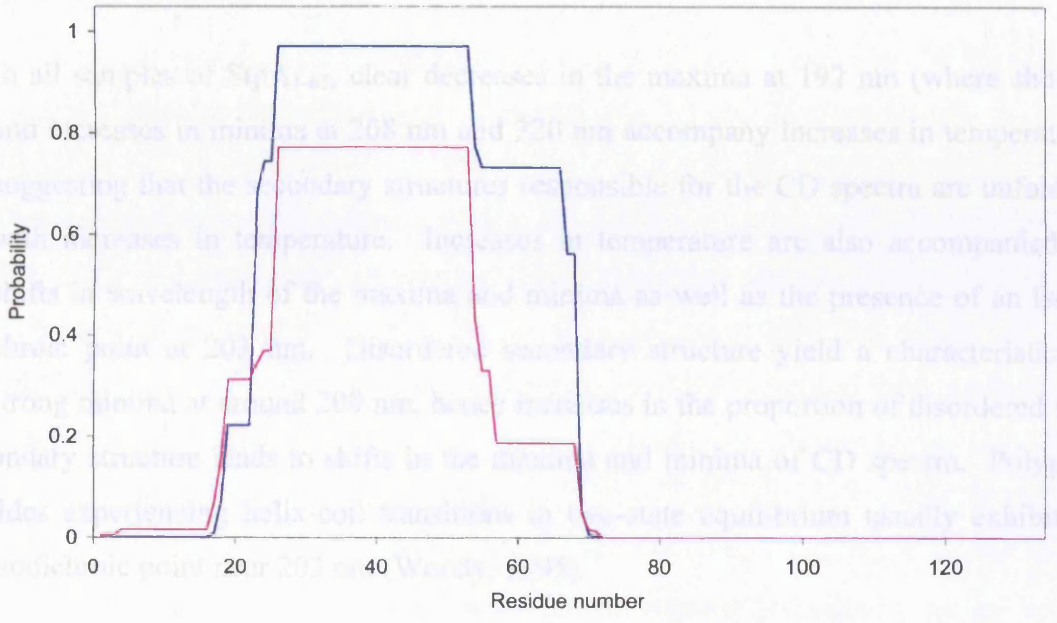


Figure 4.7
Prediction of coiled-coil formation of StpA. The probability of coiled-coil formation is plotted against StpA amino acid sequence, using either the MTIDK (—) or the MTK (—) scoring matrices. The figures were generated using Coils version 2.1, using a 28 residue scanning window, without weighting (Lupas, 1997).

4.3.5 The thermal stability of the N-terminal domain of StpA depends on salt concentration

The thermal denaturation of $StpA_{1-65}$ in a variety of salt conditions was investigated with CD spectroscopy. 30 μM samples of $StpA_{1-65}$ were prepared in a buffer of 20 mM sodium phosphate at pH 7.0 with five different concentrations of NaCl: 10mM, 50mM, 100mM, 300mM and 500mM. CD Spectra of each of these samples were recorded after increasing temperatures in increments of 10°C (between 5°C to 95°C) between 180 nm and 260 nm. Once a CD spectrum of a given sample was recorded at 95°C, another CD spectrum was recorded having lowered the temperature to 25°C in order to determine the reversibility of change in secondary structure. The $StpA_{1-65}$ samples in 500 mM NaCl, 300 mM NaCl, 100 mM NaCl, 50 mM NaCl and 10 mM NaCl are shown in Figures 4.8, 4.9, 4.10, 4.11 and 4.12 respectively. The inset diagrams in all figures demonstrate the reversibility of the thermal transitions at the respective concentrations of NaCl: sequential scans at 25°C, at 95°C and another scan at 25° taken (after the sample has cooled down) are displayed.

In all samples of $StpA_{1-65}$, clear decreases in the maxima at 192 nm (where shown) and increases in minima at 208 nm and 220 nm accompany increases in temperature, suggesting that the secondary structures responsible for the CD spectra are unfolding with increases in temperature. Increases in temperature are also accompanied by shifts in wavelength of the maxima and minima as well as the presence of an isodichroic point at 203 nm. Disordered secondary structure yield a characteristically strong minima at around 200 nm, hence increases in the proportion of disordered secondary structure leads to shifts in the maxima and minima of CD spectra. Polypeptides experiencing helix-coil transitions in two-state equilibrium usually exhibit an isodichroic point near 203 nm (Woody, 1995).

The $StpA_{1-65}$ samples in 100 mM, 300 mM and 500mM NaCl all display reversible folding. Whilst incremental increases in temperature result in gradual loss of secondary structure in the three samples, small proportions of α-helical structure remain at temperatures as high as 95°C. The spectra obtained after renaturation of $StpA_{1-65}$ (in other words, the CD scan obtained after the samples have cooled to 25°C) suggests the thermally-induced transitions of $StpA_{1-65}$ in either 100 mM, 300 mM or 500mM NaCl are reversible. In the 100mM and 300mM NaCl $StpA_{1-65}$ sample, the CD spec-

tra taken before and after thermal denaturation are slightly different, suggesting that $StpA_{1-65}$ secondary structure is not entirely recovered after renaturation in 100mM or 300mM NaCl.

In the $StpA_{1-65}$ samples containing either 10mM or 50mM NaCl, incremental increases in temperature from 5°C to 75°C result in a gradual loss of secondary structure, although a proportion of α-helical structure is retained at 75°C. However, at 85°C, the CD spectra of both samples change significantly, resulting in broad negative profiles with only one minimum. For the $StpA_{1-65}$ containing 10 mM NaCl, the minimum of the spectrum is around 215 nm, whilst the minimum of the CD spectrum corresponding to the $StpA_{1-65}$ sample containing 50 mM NaCl is found around 205 nm. β-sheet secondary structure results in a broad negative CD spectrum with a characteristic minimum at 217 nm (Woody, 1995). Therefore it is highly probable that $StpA_{1-65}$ in 10 mM NaCl at 85°C and above (and to a lesser extent, $StpA_{1-65}$ in 50 mM NaCl) exhibits significant changes in secondary structure, resulting in β-sheet-rich structures. Also the CD spectra of these two $StpA_{1-65}$ samples obtained at 85°C and above do not pass through the isodichroic point at 203 nm, suggesting the $StpA_{1-65}$ no longer exhibits the two-state helix-coil transition in the specified conditions.

The thermally induced transitions exhibits by the $StpA_{1-65}$ samples in 10 mM and 50 mM NaCl are apparently irreversible: when these samples are cooled down after thermal denaturation at 95°C and CD spectra reacquired at 25°C, the original α-helix-rich CD profile is not restored. The CD spectrum of the renatured $StpA_{1-65}$ sample at 10 mM suggests that a significant proportion of the secondary structure is β-sheet, shown by the broad minimum at 215 nm.

Polypeptides that exhibit helix-coil transitions are frequently monitored by the change in molar ellipticity at 220 nm (Woody, 1995). Molar ellipticity at 220 nm (θ_{220nm}) of the $StpA_{1-65}$ samples displaying reversible folding (i.e. the $StpA_{1-65}$ samples in 100 mM, 300 mM and 500 mM) are plotted with respect to temperature, and shown in Figure 4.13. As the presence of the isodichroic point at 203 nm implies $StpA_{1-65}$ exhibits a helix-coil transition in a two-state equilibrium, the thermally-induced changes in θ_{220nm} were analysed with sigmoidal regression. The sigmoidal function is typically used to describe two-state processes (see Section 2.12). The respective regres-

sion analyses of θ_{220nm} are plotted in Figure 4.13, with the fitting parameters displayed in Table 4.1 According to the regression analyses, the midpoint of the thermally-induced transitions (T_m) of the $StpA_{1-65}$ samples in 100 mM, 300 mM and 500 mM NaCl are 65.7°C, 72.8°C and 68.7°C, respectively. The variation of T_m between the three samples is likely to be due to experimental error; therefore the average of the three values of 69°C best describes the T_m of $StpA_{1-65}$. This T_m of $StpA_{1-65}$ is higher than that of H-NS determined in other studies. Temperature-dependent CD measurements suggested that H-NS undergoes reversible partial denaturation between 24°C and 64°C, with a T_m of 41°C (Schroder *et al.*, 2001). NMR experiments involving thermal denaturation suggested a T_m of 58 °C for full-length H-NS (Smyth *et al.*, 2000). A DSC scan of $H\text{-}NS_{1-64}$ at a concentration of 2.9 mM revealed a T_m of 52.5 °C (Smyth, 1999).

Whilst $StpA_{1-65}$ secondary structure is independent of salt concentration at lower temperatures in the range 5°C to 75°C, NaCl clearly has a role in stabilising $StpA_{1-65}$ secondary structure at temperatures higher than 75°C. $StpA_{1-65}$ displays reversible thermal denaturation in the range 5°C to 95°C at concentrations of NaCl of 100 mM or greater. However, at concentrations of NaCl of 50 mM or lower, temperatures in excess of 75°C cause an irreversible and abrupt change in secondary structure, from an α-helix-rich structure to a β-sheet-rich one. These results suggest differences in thermal stability between H-NS and StpA. H-NS was shown to exhibit reversible thermal denaturation in a buffer containing 10 mM sodium phosphate, pH 7.2 with no NaCl (Schroder *et al.*, 2001).

[NaCl] (mM)	θ_1 (deg cm^2 dmol^{-1} residue^{-1})	θ_2 (deg cm^2 dmol^{-1} residue^{-1})	T_m (°C)
100	-19848.6 ± 191.6	-15118.9 ± 605.9	65.7 ± 0.7
300	-19568.9 ± 196.4	-13862.7 ± 754.3	72.8 ± 1.0
500	-18688.5 ± 196.8	-13661.7 ± 690.3	68.7 ± 1.0

Table 4.1
The parameters of the sigmoidal regression analyses (see Section 2.12) of the change in CD signal at 220nm with respect to temperature of $StpA_{1-65}$ samples in 100 mM, 300 mM and 500 mM NaCl.

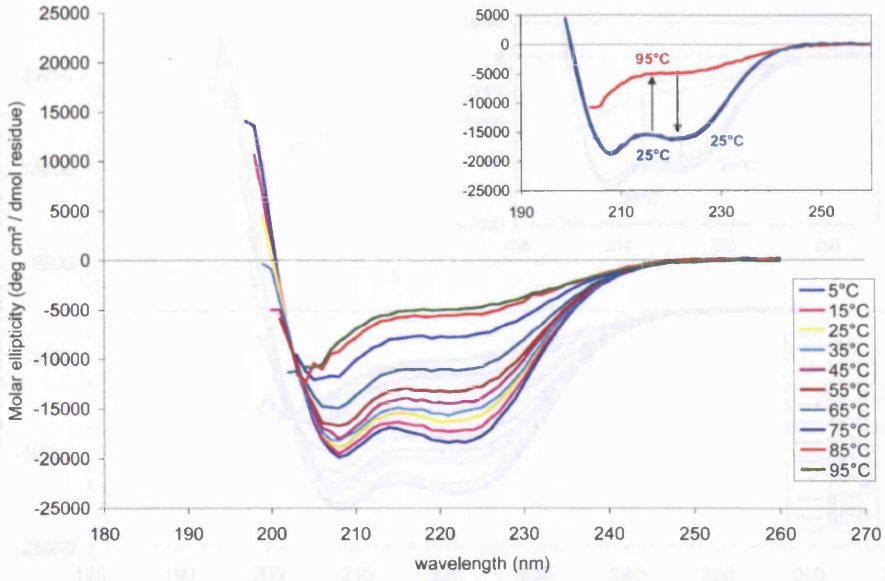


Figure 4.8
CD spectra of $StpA_{1-65}$ at 500mM NaCl at different temperatures. CD spectra were recorded on a 30 µM sample of $StpA_{1-65}$ in 500mM NaCl in a 0.1 cm path length cell in increments of 10°C. CD spectra of the different temperatures are colour-coded as per the legend above. The inset shows data selected from the same experiment: a CD wavelength scan was recorded at 25°C (—), 95°C (—), and 25°C again after cooling the sample (—).

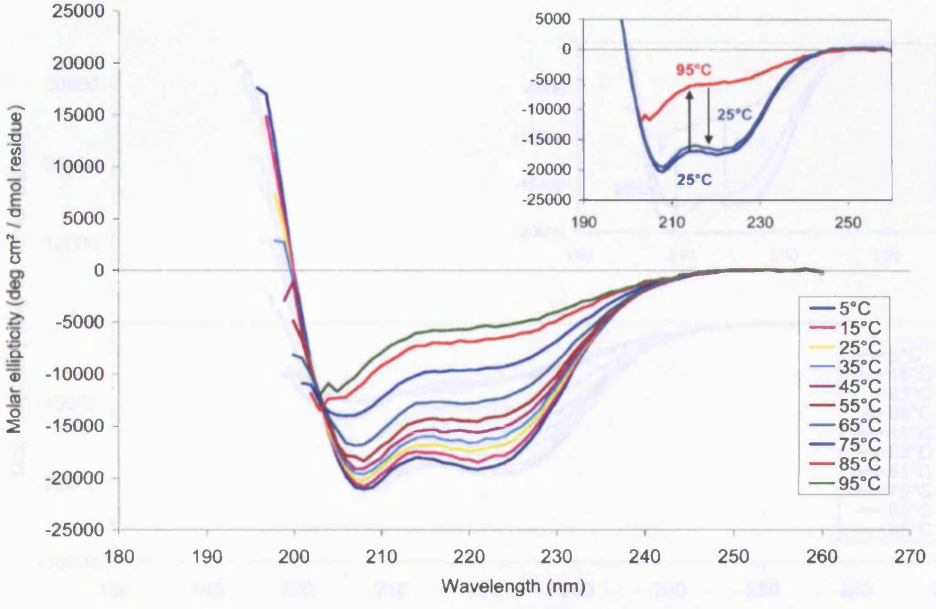


Figure 4.9
CD spectra of $StpA_{1-65}$ at 300mM NaCl at different temperatures. CD spectra were recorded on a 30 µM sample of $StpA_{1-65}$ in 300mM NaCl in a 0.1 cm path length cell in increments of 10°C. The CD spectra of the different temperatures are colour-coded as per the legend above. The inset shows data selected from the same experiment: a CD wavelength scan was recorded at 25°C (—), 95°C (—), and 25°C again after cooling the sample (—).

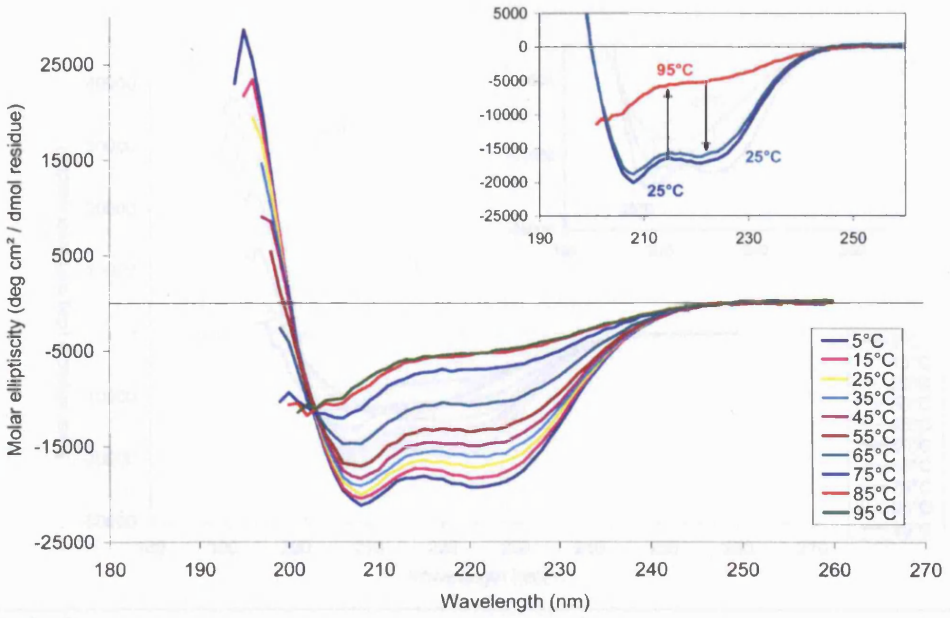


Figure 4.10
CD spectra of $StpA_{1-65}$ at 100mM NaCl at different temperatures. CD spectra were recorded on a 30 μM sample of $StpA_{1-65}$ in 100mM NaCl in a 0.1 cm path length cell in increments of 10°C. The CD spectra of the different temperatures are colour-coded as per the legend above. The inset shows data selected from the same experiment: a CD wavelength scan was recorded at 25°C (—), 95°C (—), and 25°C again after cooling the sample (—).

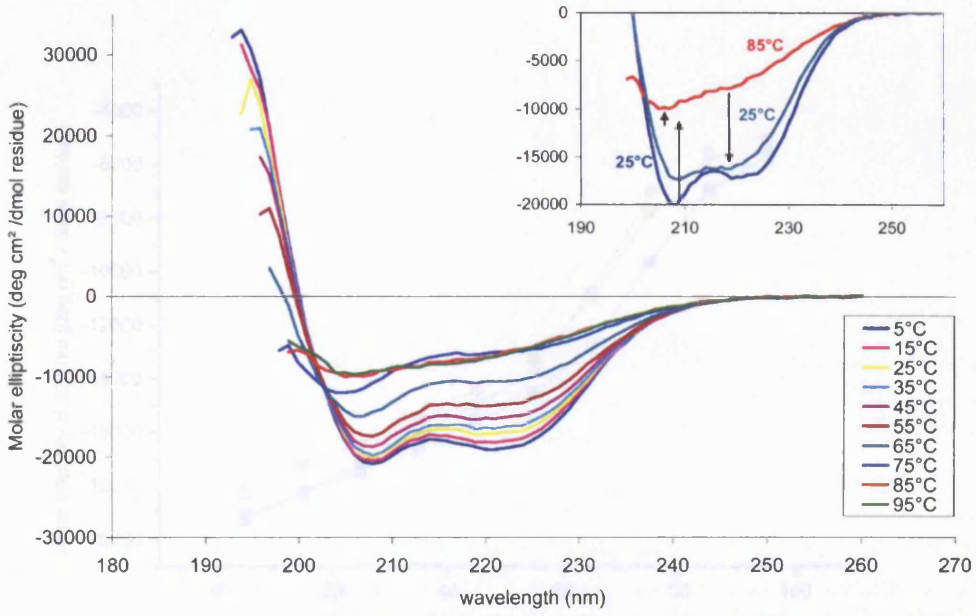


Figure 4.11
CD spectra of $StpA_{1-65}$ at 50mM NaCl at different temperatures. CD spectra were recorded on a 30 μM sample of $StpA_{1-65}$ in 50mM NaCl in a 0.1 cm path length cell in increments of 10°C. The CD spectra of the different temperatures are colour-coded as per the legend above. The inset shows data selected from the same experiment: a CD wavelength scan was recorded at 25°C (—), 95°C (—), and 25°C again after cooling the sample (—).

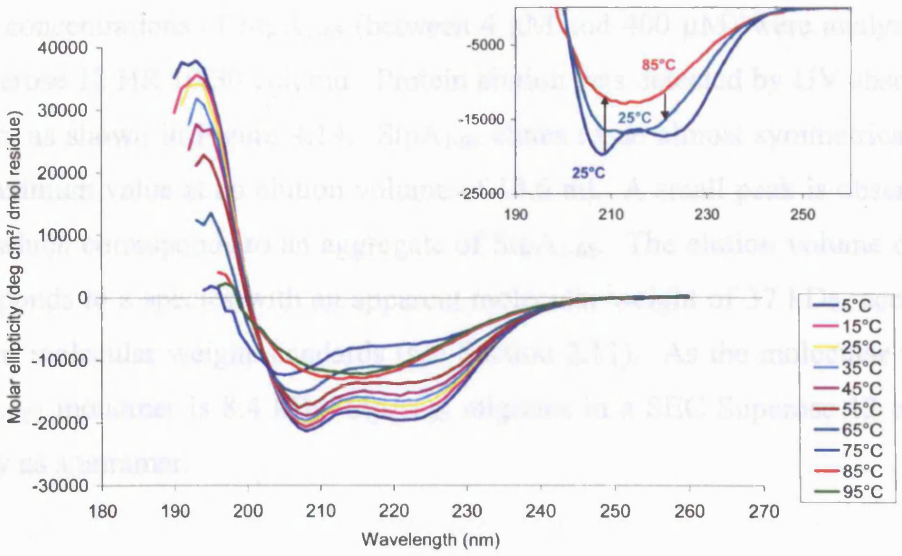


Figure 4.12
CD spectra of $StpA_{1-65}$ at 10mM NaCl at different temperatures. CD spectra were recorded on a 30 μM sample of $StpA_{1-65}$ in 10mM NaCl in a 0.1 cm path length cell in increments of 10°C. The CD spectra of the different temperatures are colour-coded as per the legend above. The inset shows data selected from the same experiment: a CD wavelength scan was recorded at 25°C (—), 95°C (—), and 25°C again after cooling the sample (—).

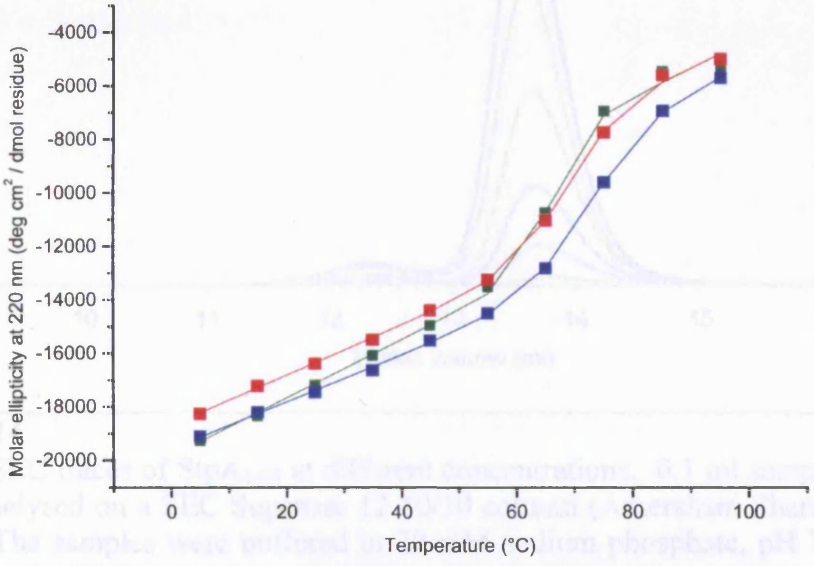


Figure 4.13
Change in molar ellipticity at 220 nm (θ_{220nm}) of the $StpA_{1-65}$ samples in 100 mM (■), 300 mM (■) and 500 mM (■) NaCl, plotted with respect to temperature. The sigmoidal curve of best fit to the θ_{220nm} data is shown on the respective figures. Note that the sample size for all data points is one.

4.3.6 The N-terminal domain of StpA forms a discrete domain

Differing concentrations of $StpA_{1-65}$ (between 4 μM and 400 μM) were analysed using a Superose 12 HR 10/30 column. Protein elution was detected by UV absorption at 280 nm, as shown in Figure 4.14. $StpA_{1-65}$ elutes as an almost symmetrical peak with a maximum value at an elution volume of 13.6 ml. A small peak is observed at 12.4 ml, which corresponds to an aggregate of $StpA_{1-65}$. The elution volume of 13.6 ml corresponds to a species with an apparent molecular weight of 37 kDa, according to globular molecular weight standards (see Section 2.11). As the molecular weight of a $StpA_{1-65}$ monomer is 8.4 kDa, $StpA_{1-65}$ migrates in a SEC Superose 12 column apparently as a tetramer.

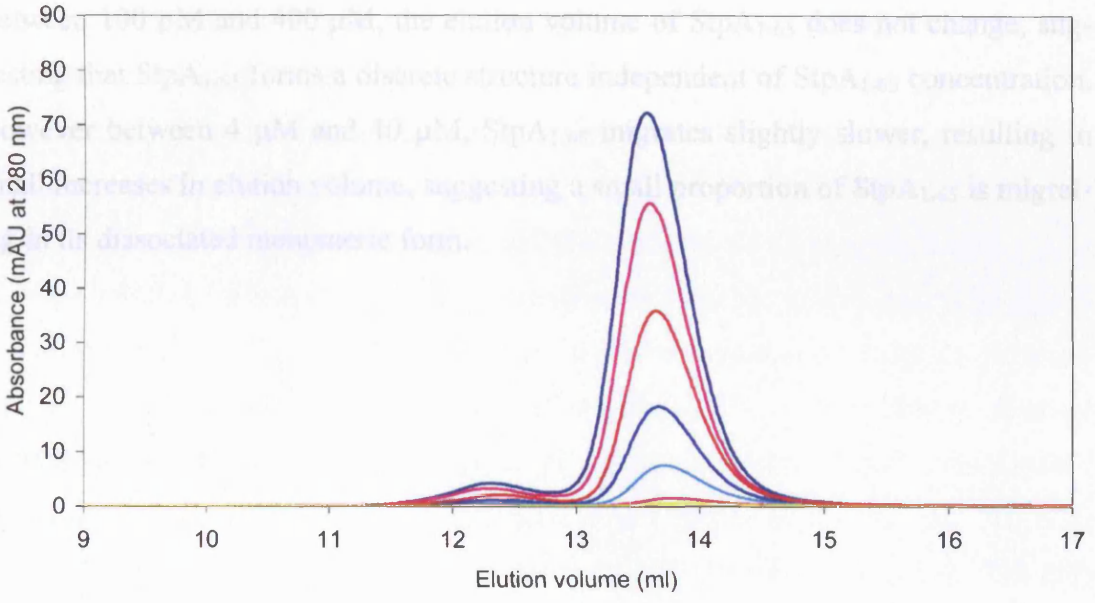


Figure 4.14
Overlaid SEC traces of $StpA_{1-65}$ at different concentrations. 0.1 ml samples of $StpA_{1-65}$ were analysed on a SEC Superose 12 10/30 column (Amersham Pharmacia) at 0.4 ml/min. The samples were buffered in 20 mM sodium phosphate, pH 7.0, 500 mM NaCl, 0.1 mM EDTA and 0.1 mM NaN_3. The following concentrations of $StpA_{1-65}$ are displayed: 400 μM (—), 300 μM (—), 200 μM (—), 100 μM (—), 40 μM (—), 10 μM (—), and 4 μM (—).

Previous SEC studies of the N-terminal domain of H-NS (H-NS_{1-64}) showed that H-NS_{1-64} migrated in an analytical SEC Superose 12 column with a velocity corresponding to a complex sized between a trimer and tetramer of H-NS_{1-64}, according to globular molecular weight standards (Smyth *et al.*, 2000). Subsequent NMR studies revealed that the N-terminal domain of H-NS (either containing residues 1 to 57 or residues 1 to 64) exists as a dimer and has an elongated structure, referred to as a 'prolate ellipsoid' (Renzoni *et al.*, 2001; Esposito *et al.*, 2002). The difference in shape of H-NS_{1-64} compared to the roughly spherical globular proteins used in SEC molecular weight standards result in discrepancies in the SEC migration behaviour between H-NS_{1-64} and the molecular weight standards, accounting for the observation that H-NS_{1-64} migrates in a SEC Superose 12 column with an apparent molecular weight almost twice the accepted value. Therefore comparison of $StpA_{1-65}$ with globular molecular weight standards is unsuitable, assuming $StpA_{1-65}$ and H-NS_{1-64} have similar elongated structures.

Between 100 μM and 400 μM, the elution volume of $StpA_{1-65}$ does not change, suggesting that $StpA_{1-65}$ forms a discrete structure independent of $StpA_{1-65}$ concentration. However between 4 μM and 40 μM, $StpA_{1-65}$ migrates slightly slower, resulting in small increases in elution volume, suggesting a small proportion of $StpA_{1-65}$ is migrating in its dissociated monomeric form.

4.4 Investigation of the interdimeric exchange of H-NS and StpA monomers

4.4.1 The monomers of the N-terminal domains of H-NS exchange slowly in solution

A study of H-NS mutants suggested that their dominant-negative activity may be a consequence of the exchange of wild-type monomers with mutant H-NS in the coiled-coil, forming a non-functional H-NS dimer (Ueguchi *et al.*, 1997). CD spectra of H-NS_{1-64} and H-NS_{FL} at differing concentrations reveal that the α-helix-rich secondary structure is not fully formed below concentrations of 30 μM. At 0.3 μM and 3 μM, the CD spectra suggest that these proteins show only a fraction of secondary structure observed in the respective fully folded proteins (Smyth *et al.*, 2000). Due to the non-covalent interaction of the monomers in the coiled-coil, the formation of the dimer is not irreversible.

Previous studies involving thermal denaturation (between 15°C and 95°C) of H-NS_{FL} with the differential scanning calorimeter (DSC) suggest that H-NS_{FL} undergoes some structural transition at around 55°C. The change of enthalpy of this transition was measured as 3.8 cal g^{-1}. The enthalpy of melting of lysozyme, a typical globular protein known to fully denature under the same conditions, was measured as 6.5 cal g^{-1}. This suggests that H-NS_{FL} has not denatured fully under these conditions, i.e. there is some degree of structure remaining in H-NS_{FL} even at 95°C (Smyth, 1999). This provides indirect evidence of the high thermal stability of the coiled-coil. The transition observed at 55°C during the DSC scan of H-NS_{FL} may be attributed to the melting of C-terminal DNA-binding domain of H-NS_{FL} and/or to the melting of some structure surrounding the coiled-coil in the N-terminal domain. Other thermal stability studies include tracking temperature-dependent changes in H-NS_{FL} and H-NS_{1-64} with CD spectroscopy, where CD spectra (measured between 190 nm to 260 nm) were obtained at a range of temperatures up to 82°C (Smyth, 1999). Both constructs displayed CD spectra consistent with α-helical-rich secondary structure. Both structures retained a certain degree of α-helical secondary structure at temperatures as high as 82°C, suggesting H-NS has a high thermal stability in some or all of its structure, most likely due to the coiled-coil structure.

However no study to date has investigated the degree of exchange of the monomeric units between N-terminal coiled-coil dimers. 'Pull-down' experiments were per-

formed to provide qualitative data concerning the degree to which the monomers in the coiled-coil of H-NS exchange. These experiments allow isolation of H-NS_{1-64} that have exchanged the monomers of the coiled-coil region (in other words, H-NS_{1-64} dimers that had undergone exchange are 'pulled down' from solution). The experimental details are described in Section 2.15. The results of the 'pull-down' assay were analysed by 12 % SDS-PAGE. The samples stored at 4°C are shown in Figure 4.15 and the samples stored at 37°C are shown in Figure 4.16.

A control sample containing only 100 μM His-tagged H-NS_{1-64} was applied to the TALON column as described in Section 2.15. In both Figures 4.15 and 4.16, absence of protein in the lane 1 flow through sample and presence of His-tagged H-NS_{1-64} in the lane 2 elution sample clearly show that His-tagged H-NS_{1-64} binds to the TALON column well. Another control sample containing only 100 μM untagged H-NS_{1-64} was applied to the TALON column as described. Presence of H-NS_{1-64} in the lane 3 flow-through sample and absence of protein in the lane 4 elution sample show that without a His-tag, H-NS_{1-64} fails to bind to the TALON column. In lane 2 corresponding to His-tagged H-NS_{1-64} (molecular weight 9.9 kDa), there are three visible bands. Band A found at 20 kDa corresponds to His-tagged H-NS_{1-64} dimer that has not been denatured fully by SDS. Band B found at 10 kDa corresponds to the monomeric His-tagged H-NS_{1-64}. Band C is a small band that is likely to correspond to degradation products, and can be ignored. In Lane 3, corresponding to H-NS_{1-64} (molecular weight of 8.0 kDa), there are two visible bands. Band D found at 16 kDa corresponds to dimeric H-NS_{1-64} that has not been fully denatured by SDS. Band E found at 8 kDa corresponds to monomeric H-NS_{1-64}.

Lanes 8 and 9 of Figures 4.15 and 4.16 show the flow-through and elution of the refolded sample (treated with 6 M guanidine hydrochloride at 55°C for one hour). Upon denaturation and refolding, an equimolar mixture of His-tagged H-NS_{1-64} and untagged H-NS_{1-64} should contain heterodimers of His-tagged-H-NS_{1-64}/untagged H-NS_{1-64}, as well as homodimers of His-tagged H-NS_{1-64} and homodimers of untagged H-NS_{1-64}: the refolded sample represents a control in which there has been complete exchange of the monomeric subunits between the coiled-coil dimers. The column flow-through shown in lane 8 contains only H-NS_{1-64} and no His-tagged H-NS_{1-64}, as expected. The elution sample, shown in lane 9, displays several bands: band F corre-

sponds to dimeric His-tagged H-NS_{1-64} and band G at around 17 kDa corresponds to a heterodimer consisting of His-tagged H-NS_{1-64} and untagged H-NS_{1-64}. As expected, dimeric untagged H-NS_{1-64} is not observed, as it would not be expected to bind the column. Band H corresponds to monomeric His-tagged H-NS_{1-64}. Band I corresponds to degradation products (equivalent to band C in lane 2). Band J corresponds to monomeric untagged H-NS_{1-64}. The presence of the heterodimers of His-tagged-H-NS_{1-64}/untagged H-NS_{1-64} in band G and the presence of untagged H-NS_{1-64} monomer in band J indicate that exchange of the coiled-coil has occurred. In other words, H-NS_{1-64} cannot bind the column as a dimer, and therefore will not be present in the elution: on the other hand, the heterodimers of His-tagged-H-NS_{1-64}/untagged H-NS_{1-64} is able to bind the column, resulting in the visible bands of untagged H-NS_{1-64} in the elution.

Lanes 5, 6 and 7 of Figure 4.15 correspond to samples that have been stored at 4°C for 1 hour, 15 hours and 40 hours respectively. All samples show bands corresponding to the presence of untagged H-NS_{1-64}, i.e. the equivalent to band J found in the refolded sample. However none of the lanes show the band corresponding to the heterodimer consisting of His-tagged H-NS_{1-64}/untagged H-NS_{1-64}, i.e. the equivalent to band G found in the refolded sample. This indicates that there has been a small degree of exchange of the coiled-coil in the conditions specified.

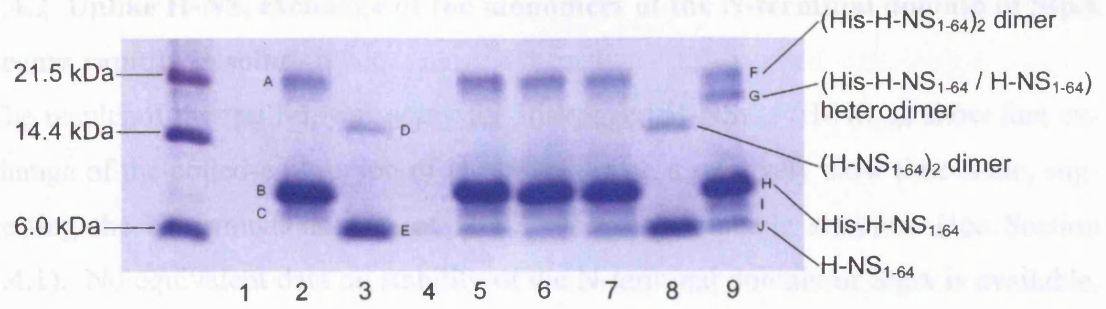
Lanes 5, 6 and 7 of Figure 4.16 correspond to samples that have been stored at 37°C for 1 hour, 15 hours and 40 hours respectively. After 1 hour at 37°C, only a small amount of exchange between homodimers of untagged H-NS_{1-64} and His-tagged H-NS_{1-64} has occurred, as shown by the presence of a small amount of untagged H-NS_{1-64}, equivalent to band J of the refolded sample. Only a small amount of the heterodimer consisting of His-tagged H-NS_{1-64}/H-NS_{1-64} is observed, i.e. equivalent to band G in the refolded sample. After 15 or 40 hours at 37°C, there has been complete exchange, i.e. the exchange between the untagged H-NS_{1-64} and His-tagged H-NS_{1-64} has reached equilibrium, as shown by the presence of bands equivalent to bands G and J of the refolded sample.

Results from the co-incubation of untagged H-NS_{1-64} and His-tagged H-NS_{1-64} show that exchange occurs between the monomers of the coiled-coils of the respective H-

NS_{1-64} dimers. This exchange is a slow process, indicating the high stability afforded by the formation of the coiled-coil: exchange between the coiled-coils at 4°C has not reached equilibrium even after 40 hours of incubation. At 37°C, the apparent rate of exchange is increased such that equilibrium is reached within 15 hours. This proposal of the high stability of the H-NS N-terminal dimer is consistent with the DSC data (Smyth, 1999).

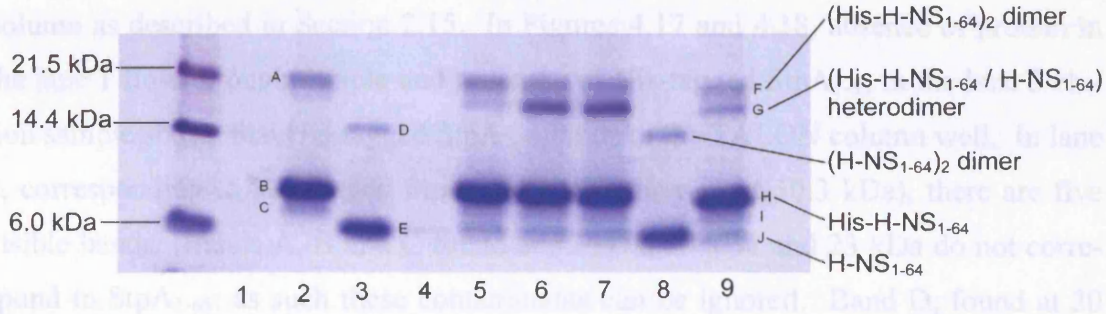
However these conclusions should be accepted with caution, as a significant weight has been attributed in the interpretation of the data to the differential levels of the heterodimer consisting of His-tagged H-NS_{1-64}/H-NS_{1-64}. Control experiments were not conducted to determine whether the complexes that resulted in the corresponding bands in the SDS-PAGE gels were native or non-specific aggregates.

If the bands in the SDS-PAGE gels corresponding to heterodimers of His-tagged H-NS_{1-64}/untagged H-NS_{1-64} (i.e. band G in Figure 4.15 and 4.16) are indeed from non-specific aggregates, they must be excluded from interpretation. Thus all samples, stored at both temperatures of 4°C and 37°C, display comparable amounts of 'pulled-down' untagged H-NS_{1-64} with respect to the control samples, suggesting that the monomers of H-NS N-terminal domains exchange rapidly in solution. However it is worth noting that freshly prepared H-NS_{1-64} still yields a band in SDS-PAGE gels corresponding to the dimer (data not shown). This suggests that the coiled-coil affords the H-NS dimer with high stability, and is not completely denatured under the conditions of SDS-PAGE. Thus the band G observed in the various samples shown in Figures 4.15 and 4.16 may not necessarily represent non-specific aggregates.



1 His-tagged H-NS_{1-64}, flow-through
2 His-tagged H-NS_{1-64}, elution
3 Untagged H-NS_{1-64}, flow-through
4 Untagged H-NS_{1-64}, elution
5 1 hour at 4°C, elution
6 15 hours at 4°C, elution
7 40 hours at 4°C, elution
8 Refolded sample, flow-through
9 Refolded sample, elution

Figure 4.15
His-tag pull-down assay of equimolar mixtures of His-tagged H-NS_{1-64} and untagged H-NS_{1-64} at 100 μM, incubated for either 1, 15 or 40 hours at 4°C. TALON spin column flow-through and elution samples were analysed by 12% SDS-PAGE, with selected samples shown above. Lanes 1-9 were loaded with samples according to the table above. Letters A-J are used to annotate selected bands, and correspond to a description of the diagram in Section 4.4.1. A selection of the bands has been labelled to aid identification.



1 His-tagged H-NS_{1-64}, flow-through
2 His-tagged H-NS_{1-64}, elution
3 Untagged H-NS_{1-64}, flow-through
4 Untagged H-NS_{1-64}, elution
5 1 hour at 37°C, elution
6 15 hours at 37°C, elution
7 40 hours at 37°C, elution
8 Refolded sample, flow-through
9 Refolded sample, elution

Figure 4.16
His-tag pull-down assay of a mixture of His-tagged H-NS_{1-64} and untagged H-NS_{1-64}, incubated for either 1, 15 or 40 hours at 37°C. TALON spin column flow-through and elution samples were analysed by 12% SDS-PAGE, with selected samples shown above. Lanes 1-9 were loaded with samples according to the table above. Letters A-J are used to annotate selected bands, and correspond to a description of the diagram in Section 4.4.1. A selection of the bands has been labelled to aid identification.

4.4.2 Unlike H-NS, exchange of the monomers of the N-terminal domain of StpA occurs rapidly in solution

The results of the 'pull-down' assay for His-tagged H-NS_{1-64} /H-NS_{1-64} show that exchange of the coiled-coil region of H-NS occurs on a relatively slow time-scale, suggesting the N-terminal domain of H-NS represents a stable structure (see Section 4.4.1). No equivalent data on stability of the N-terminal domain of StpA is available. CD studies, described in Section 4.3.5, suggest that the thermal stability of the N-terminal domain of StpA is different than that of H-NS, at least in low salt conditions.

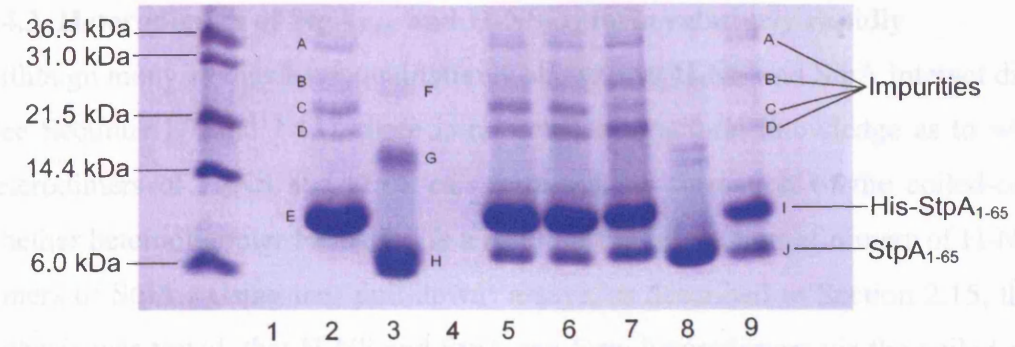
A 'pull-down' assay, designed to isolate $StpA_{1-65}$ species that had undergone exchange, was carried out to qualitatively determine the stability of the N-terminal domain of StpA, as described in Section 2.15, using mixtures of 100 μM His-tagged $StpA_{1-65}$ and 100 μM untagged $StpA_{1-65}$, at two temperatures, 4°C and 37°C. The samples were analysed by 12 % SDS-PAGE; the gel containing the samples incubated at 4°C are shown in Figure 4.17, and the gel containing the samples incubated at 37°C are shown in Figure 4.18.

A control containing only 100 μM His-tagged $StpA_{1-65}$ was applied to the TALON column as described in Section 2.15. In Figures 4.17 and 4.18, absence of protein in the lane 1 flow through sample and presence of His-tagged $StpA_{1-65}$ in the lane 2 elution sample shows that His-tagged $StpA_{1-65}$ binds to the TALON column well. In lane 2, corresponding to His-tagged $StpA_{1-65}$ (molecular weight 10.3 kDa), there are five visible bands. Bands A, B and C found at 35 kDa, 30 kDa and 23 kDa do not correspond to $StpA_{1-65}$: as such these contaminants can be ignored. Band D, found at 20 kDa, is likely to correspond to dimeric His-tagged $StpA_{1-65}$. Band E found at 10 kDa corresponds to monomeric His-tagged $StpA_{1-65}$. Another control containing only 100 μM untagged $StpA_{1-65}$ was applied to the TALON column as described. Presence of untagged $StpA_{1-65}$ in the lane 3 flow-through sample and absence of protein in the lane 4 elution sample shows that untagged $StpA_{1-65}$ fails to bind to the TALON column due to the lack of a His-tag. In lane 3, corresponding to $StpA_{1-65}$ (molecular weight of 8.4 kDa), there are three visible bands. Band F found at 25 kDa is likely to correspond to an aggregate of three $StpA_{1-65}$ monomers, and represents an impurity that can be ignored. Band G corresponds to dimeric untagged $StpA_{1-65}$. Band H, found at 8 kDa, corresponds to monomeric $StpA_{1-65}$.

Lanes 8 and 9 in Figures 4.17 and 4.18 show the flow-through and elution of the refolded sample of His-tagged $StpA_{1-65}$ and untagged $StpA_{1-65}$ (treated with 6 M guanidine hydrochloride at 55°C for one hour). The refolded sample represents a control in which there has been complete exchange of the monomeric subunits between the coiled-coil dimers. The column flow-through shown in lane 8 contains only untagged $StpA_{1-65}$ and no His-tagged $StpA_{1-65}$, as expected. The elution sample, shown in lane 9, displays several bands: however only bands I and J are of interest. Band I consists of monomeric His-tagged $StpA_{1-65}$ and band J consists of monomeric $StpA_{1-65}$. The other bands are likely to constitute impurities. The presence of the untagged $StpA_{1-65}$ monomer (corresponding to band J) indicates that exchange of the coiled-coil has occurred. Although untagged $StpA_{1-65}$ cannot interact with the TALON column by itself, by forming a heterodimer of His-tagged $StpA_{1-65}$ /untagged $StpA_{1-65}$ via exchange of the coiled-coil domain, untagged $StpA_{1-65}$ may bind the TALON column indirectly.

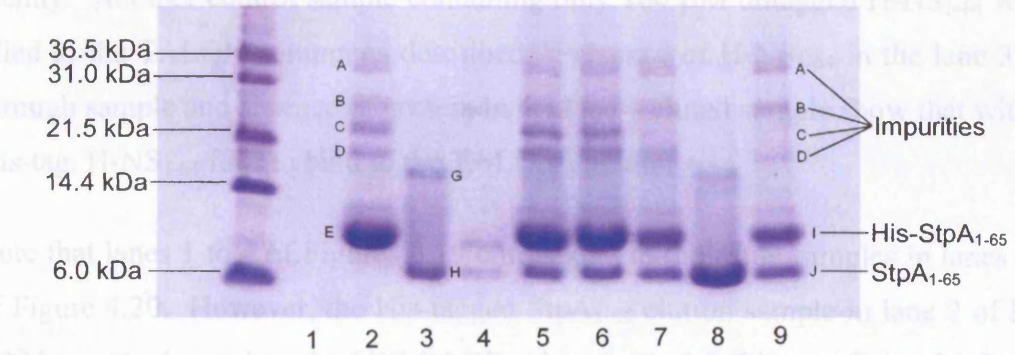
Lane 5, 6 and 7 in Figure 4.17 correspond to sample mixtures of His-tagged $StpA_{1-65}$ and untagged $StpA_{1-65}$ that have been stored at 4°C for 1 hour, 15 hours and 40 hours respectively. All samples show bands corresponding to the presence of untagged $StpA_{1-65}$, i.e. the equivalent to band J found in the refolded sample. Note that untagged $StpA_{1-65}$ is detectable in the elution samples even after 1 hour at 4°C (in amounts comparable to the refolded sample), suggesting that the exchange of the coiled-coils of $StpA_{1-65}$ have fully exchanged. Unfortunately, due to the large number of impurities at molecular weights greater than 20 kDa, these higher molecular weight bands cannot be included in interpretation of the data. Lane 5, 6 and 7 in Figure 4.18 correspond to samples of His-tagged $StpA_{1-65}$ and untagged $StpA_{1-65}$ that have been stored at 37°C for 1 hour, 15 hours and 40 hours respectively. Like the samples incubated at 4°C, all samples show presence of untagged $StpA_{1-65}$, i.e. the equivalent to band J found in the refolded sample, in quantities comparable to the refolded sample.

The results from Figures 4.17 and 4.18 suggest that complete exchange of the monomers of the coiled-coils of the His-tagged $StpA_{1-65}$ and untagged $StpA_{1-65}$ occurs in less than one hour, at either 4°C or 37°C. In comparison to H-NS, the exchange of the N-terminal domain of StpA is relatively rapid, perhaps indicating a lower stability of the $StpA_{1-65}$ N-terminal domain compared to H-NS.



1	His-tagged $StpA_{1-65}$, flow-through	5	1 hour at 4°C, elution
2	His-tagged $StpA_{1-65}$, elution	6	15 hours at 4°C, elution
3	Untagged $StpA_{1-65}$, flow-through	7	40 hours at 4°C, elution
4	Untagged $StpA_{1-65}$, elution	8	Refolded sample, flow-through
		9	Refolded sample, elution

Figure 4.17
His-tag pull-down assay of a mixture of His-tagged $StpA_{1-65}$ and untagged $StpA_{1-65}$, incubated for either 1, 15 or 40 hours at 4°C. TALON spin column flow-through and elution samples were analysed by 12% SDS-PAGE, with selected samples shown above. Lanes 1-9 were loaded with samples according to the table above. Letters A-J are used to annotate selected bands, and correspond to a description of the diagram in Section 4.4.2. A selection of the bands has been labelled to aid identification.



1	His-tagged $StpA_{1-65}$, flow-through	5	1 hour at 37°C, elution
2	His-tagged $StpA_{1-65}$, elution	6	15 hours at 37°C, elution
3	Untagged $StpA_{1-65}$, flow-through	7	40 hours at 37°C, elution
4	Untagged $StpA_{1-65}$, elution	8	Refolded sample, flow-through
		9	Refolded sample, elution

Figure 4.18
His-tag pull-down assay of a mixture of His-tagged $StpA_{1-65}$ and untagged $StpA_{1-65}$ incubated for either 1, 15 or 40 hours at 37°C. TALON spin column flow-through and elution samples were analysed by 12% SDS-PAGE, with selected samples shown above. Lanes 1-9 were loaded with samples according to the table above. Letters A-J are used to annotate selected bands, and correspond to a description of the diagram in Section 4.4.2. A selection of the bands has been labelled to aid identification.

4.4.3 Heterodimers of $StpA_{1-65}$ and H-NS_{1-64} form relatively rapidly

Although many studies have qualitatively shown that H-NS and StpA interact directly (see Sections 1.7 and 1.13), there is no detailed structural knowledge as to whether heterodimers of H-NS and StpA can form via the formation of the coiled-coil, or whether heterooligomer formation is a result of the interaction of dimers of H-NS and dimers of StpA. Using the 'pull-down' assays, as described in Section 2.15, the hypothesis was tested, that H-NS and StpA can form heterodimers via the coiled-coil of the N-terminal domain. Mixtures of 100 μM His-tagged $StpA_{1-65}$ and 100 μM untagged H-NS_{1-64} were incubated at two temperatures, 4°Cand 37°C, for 1, 15 or 40 hours. The samples were analysed by SDS-PAGE; the gel containing the samples incubated at 4°C are shown in Figure 4.19, and the gel containing the samples incubated at 37°C are shown in Figure 4.20.

A control sample containing only 100 μM His-tagged $StpA_{1-65}$ was applied to the TALON column as described in Section 2.15. In Figures 4.19 and 4.20, absence of protein in the lane 1 flow through sample and presence of His-tagged $StpA_{1-65}$ in the lane 2 eluted sample show that His-tagged $StpA_{1-65}$ binds to the TALON column efficiently. Another control sample containing only 100 μM untagged H-NS_{1-64} was applied to the TALON column as described. Presence of H-NS_{1-64} in the lane 3 flow-through sample and absence of protein in the lane 4 eluted sample show that without a His-tag, H-NS_{1-64} fails to bind to the TALON column.

Note that lanes 1 to 4 of Figures 4.19 correspond to the same samples in lanes 1 to 4 of Figure 4.20. However, the His-tagged $StpA_{1-65}$ elution sample in lane 2 of Figure 4.20 has not migrated on the SDS-PAGE gel in an ideal fashion, so instead refer to the equivalent sample shown in lane 2 of Figure 4.19. In Figure 4.19, absence of protein in the lane 1 flow through sample and presence of His-tagged $StpA_{1-65}$ in the lane 2 elution sample shows that His-tagged $StpA_{1-65}$ binds to the TALON column well. In lane 2 of Figure 4.19, corresponding to His-tagged $StpA_{1-65}$ (molecular weight 10.3 kDa), there are five visible bands. Bands A, B and C found at 35 kDa, 30 kDa and 23 kDa do not correspond to $StpA_{1-65}$, and hence these contaminants may be ignored. Band D, found at 20 kDa, is likely to correspond to dimeric His-tagged $StpA_{1-65}$. Band E found at 10 kDa corresponds to monomeric His-tagged $StpA_{1-65}$.

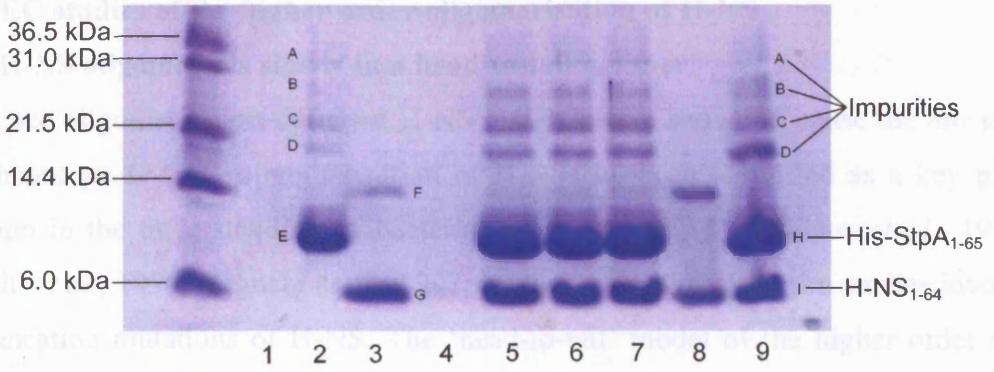
A control containing only 100 μM untagged H-NS_{1-64} was applied to the TALON column as described in Section 2.15. Presence of untagged H-NS_{1-64} in the lane 3 flow-through sample and absence of protein in the lane 4 elution sample shows that untagged H-NS_{1-64} fails to bind to the TALON column due to the lack of a His-tag. In lane 3, corresponding to H-NS_{1-64} (molecular weight of 8.0 kDa), there are two visible bands. Band F corresponds to dimeric untagged H-NS_{1-64} and band G corresponds to monomeric untagged H-NS_{1-64}.

Lanes 8 and 9 in Figures 4.19 and 4.20 show the flow-through and elution of the refolded sample (treated with 6 M guanidine hydrochloride at 55°C for one hour). Upon denaturation and refolding, an equimolar mixture of His-tagged $StpA_{1-65}$ and untagged H-NS_{1-64} should contain heterodimers of His-tagged-$StpA_{1-65}$/untagged H-NS_{1-64}, as well as homodimers of His-tagged $StpA_{1-65}$ and homodimers of untagged H-NS_{1-64}: the refolded sample constitutes a control in which there has been complete exchange of the monomeric subunits between the coiled-coil dimers. The column flow-through shown in lane 8 contains only H-NS_{1-64} and no His-tagged $StpA_{1-65}$, as expected. The elution sample, shown in lane 9, displays several bands: only bands H and I are of interest, as the other bands are likely to correspond to impurities. Band H corresponds to monomeric His-tagged $StpA_{1-65}$, and band I corresponds to monomeric untagged H-NS_{1-64}. The presence of untagged H-NS_{1-64} in band I in a TALON spin column elution sample indicates that exchange of the coiled-coil has occurred. In other words, untagged H-NS_{1-64} cannot bind the column as a dimer, and therefore will not be present in the elution: on the other hand, the heterodimers of His-tagged-H-NS_{1-64}/untagged H-NS_{1-64} are able to bind the column, resulting in the visible bands of untagged H-NS_{1-64} in the elution.

Lanes 5, 6 and 7 of Figure 4.19 correspond to samples that have been stored at 4°C for 1 hour, 15 hours and 40 hours respectively. All samples show bands corresponding to the presence of untagged H-NS_{1-64}, i.e. the equivalent to band I found in the refolded sample. This indicates that there has been exchange of the coiled-coils in the conditions specified. Lane 5, 6 and 7 in Figure 4.20 correspond to samples of His-tagged $StpA_{1-65}$ and untagged H-NS_{1-64} that have been stored at 37°C for 1 hour, 15 hours and 40 hours respectively. Like the samples incubated at 4°C, all samples show

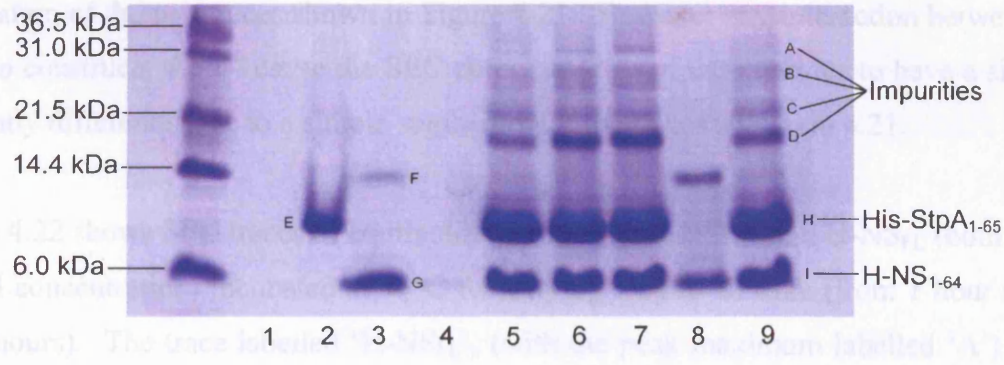
presence of untagged H-NS_{1-64}, i.e. the equivalent to band I found in the refolded sample, in quantities comparable to the refolded sample.

Results from the co-incubation of His-tagged $StpA_{1-65}$ and untagged H-NS_{1-64} suggest that the monomers exchange between the coiled-coils of the $StpA_{1-65}$ and H-NS_{1-64} dimers. The intensities of the bands corresponding to untagged H-NS_{1-64} does not vary between the samples, regardless of incubation time or temperature, suggesting that at 4°C or 37°C, complete exchange has occurred between the $StpA_{1-65}$ and H-NS_{1-64} in less than one hour, resulting in the formation of coiled-coil heterodimers of $StpA_{1-65}$ and H-NS_{1-64}.



1 His-tagged $StpA_{1-65}$, flow-through
2 His-tagged $StpA_{1-65}$, elution
3 Untagged $H\text{-}NS_{1-64}$, flow-through
4 Untagged $H\text{-}NS_{1-64}$, elution
5 1 hour at 4°C, elution
6 15 hours at 4°C, elution
7 40 hours at 4°C, elution
8 Refolded sample, flow-through
9 Refolded sample, elution

Figure 4.19
His-tag pull-down assay of a mixture of His-tagged $StpA_{1-65}$ and untagged $H\text{-}NS_{1-64}$ incubated for either 1, 15 or 40 hours at 4°C. TALON spin column flow-through and elution samples were analysed by 12% SDS-PAGE, with selected samples shown above. Lanes 1-9 were loaded with samples according to the table above. Letters A-I are used to annotate selected bands, and correspond to a description of the diagram in Section 4.4.3. A selection of the bands has been labelled to aid identification.



1 His-tagged $StpA_{1-65}$, flow-through
2 His-tagged $StpA_{1-65}$, elution
3 Untagged $H\text{-}NS_{1-64}$, flow-through
4 Untagged $H\text{-}NS_{1-64}$, elution
5 1 hour at 37°C, elution
6 15 hours at 37°C, elution
7 40 hours at 37°C, elution
8 Refolded sample, flow-through
9 Refolded sample, elution

Figure 4.20
His-tag pull-down assay of a mixture of His-tagged $StpA_{1-65}$ and untagged $H\text{-}NS_{1-64}$ incubated for either 1, 15 or 40 hours at 37°C. TALON spin column flow-through and elution samples were analysed by 12% SDS-PAGE, with selected samples shown above. Lanes 1-9 were loaded with samples according to the table above. Letters A-I are used to annotate selected bands, and correspond to a description of the diagram in Section 4.4.3. A selection of the bands has been labelled to aid identification.

4.5. SEC studies of the higher order oligomerisation of H-NS

4.5.1 H-NS oligomerises slowly in a head-to-tail manner

Ever since oligomerisation-deficient H-NS mutants were shown to cause the *hns* mutant phenotype *in vivo*, oligomerisation of H-NS has been identified as a key phenomenon in the understanding of bacterial DNA packaging (Williams *et al.*, 1996; Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997; Badaut *et al.*, 2002). From studies involving truncation mutations of H-NS, The 'head-to-tail' model of the higher order oligomerisation of H-NS has been proposed (Esposito *et al.*, 2002; Badaut *et al.*, 2002), and is discussed in detail in Section 1.9.4.

The effect of H-NS_{1-64} on the higher order oligomerisation of H-NS_{FL} was investigated by SEC. Mixtures of H-NS_{FL} and H-NS_{1-64} were subjected to a variety of storage conditions such as differing temperatures, periods of incubation and ratios of H-NS_{FL} to H-NS_{1-64}, and then analysed using SEC. Figure 4.21 shows SEC traces of H-NS_{FL} and H-NS_{1-64}. Both samples contained 40 μM of protein and were run separately. The difference in peak area merely reflects the difference in per-molar extinction coefficients of the two constructs. If the two constructs do not interact with each other, a mixture of the two samples should reveal a chromatogram that was simply the summation of the two traces shown in Figure 4.21. However any interaction between the two constructs should cause the SEC chromatogram of the mixtures to have a significantly different shape to a simple summation of the traces shown in 4.21.

Figure 4.22 shows SEC traces of equimolar mixtures of H-NS_{1-64} and H-NS_{FL} (both at 40 μM concentration) incubated at 45°C for varying lengths of time (from 1 hour up to 24 hours). The trace labelled 'H-NS_{FL}', (with the peak maximum labelled 'A') is from a sample containing only 40 μM H-NS_{FL}. Peak A corresponds to the elution of the mode molecular weight of H-NS_{FL} under these conditions. Peak C corresponds to the mode molecular weight species in the mixtures of H-NS_{FL} and H-NS_{1-64}. Peak D corresponds to H-NS_{1-64}.

The addition of H-NS_{1-64} into a solution of H-NS_{FL} significantly changes the nature of the population of oligomers, seen as a dramatic change in the SEC traces after the addition of H-NS_{1-64}. The elution volume of the mode molecular mass species found in a mixture of H-NS_{FL} and H-NS_{1-64} (corresponding to the collection of peaks labelled

'B') is significantly higher than the elution volume of the mode molecular mass species of a sample of H-NS_{FL} alone (the peak labelled 'A'). This observation implies that the oligomeric species found in a mixture of H-NS_{1-64} and H-NS_{FL} is significantly smaller than those found in a solution of H-NS_{FL} alone. As the incubation period was increased from one hour to 24 hours, the mode molecular weight found in the mixture of H-NS_{FL} and H-NS_{1-64} progressively decreased. This progressive decrease of modal oligomer size was mirrored by a progressive increase of a discrete peak found at an elution volume of 12.4 ml (labelled peak 'C'). According to a globular protein standard curve for this column, peak C represents a species with a molecular mass of 75 kDa. A heterotetramer of H-NS_{FL} and H-NS_{1-64}, with a predicted molecular weight of 47.6 kDa, may account for this peak. H-NS_{1-64} is known to migrate in the SEC column in an non-ideal fashion: the H-NS_{1-64} dimer (of molecular weight 16 kDa) has been shown to migrate at an apparent molecular weight of approximately 28 kDa (Smyth, 1999), hence a heterotetramer of H-NS_{FL} and H-NS_{1-64} may migrate at a higher apparent molecular weight than predicted. Peak C reaches it maximal value after approximately 5 hours.

According to the head-to-tail model, H-NS_{FL} retains all the residues identified as critical for higher order oligomerisation, i.e. H-NS_{FL} retains both the 'head' and the 'tail', whereas H-NS_{1-64} (lacking residues 65 to 89) retains only the 'tail'. According to the 'head-to-tail' model, H-NS_{1-64} can interact with a homodimer of H-NS_{FL}, mediated by the 'tail' of a homodimer of H-NS_{1-64} with the 'head' of the homodimer of H-NS_{FL}. This heterotetramer complex (consisting of two H-NS_{FL} molecules and two H-NS_{1-64} molecules) represents the smallest oligomeric species that can be found in a mixture of H-NS_{FL} and H-NS_{1-64}. Therefore this model predicts the formation of the following complexes in a mixture of H-NS_{FL} and H-NS_{1-64}: oligomers of H-NS_{FL} of indeterminate size, oligomers of H-NS_{FL} of indeterminate size with a dimer of H-NS_{1-64} bound to the 'head' end of the oligomer, a heterotetramer of H-NS_{FL} dimer and H-NS_{1-64} and dimer, and dimeric H-NS_{1-64}. Indeed oligomers (of higher order than a tetramer) are found in the SEC traces corresponding to the collection of peaks B, the heterodimer complex of H-NS_{FL} and H-NS_{1-64} is found as peak C, and uncomplexed H-NS_{1-64} is found as peak D. These SEC traces provide further evidence to back up the 'head-to-tail' model for the oligomerisation of H-NS, as well as giving an indication of the time-scale over which the oligomerisation processes occur.

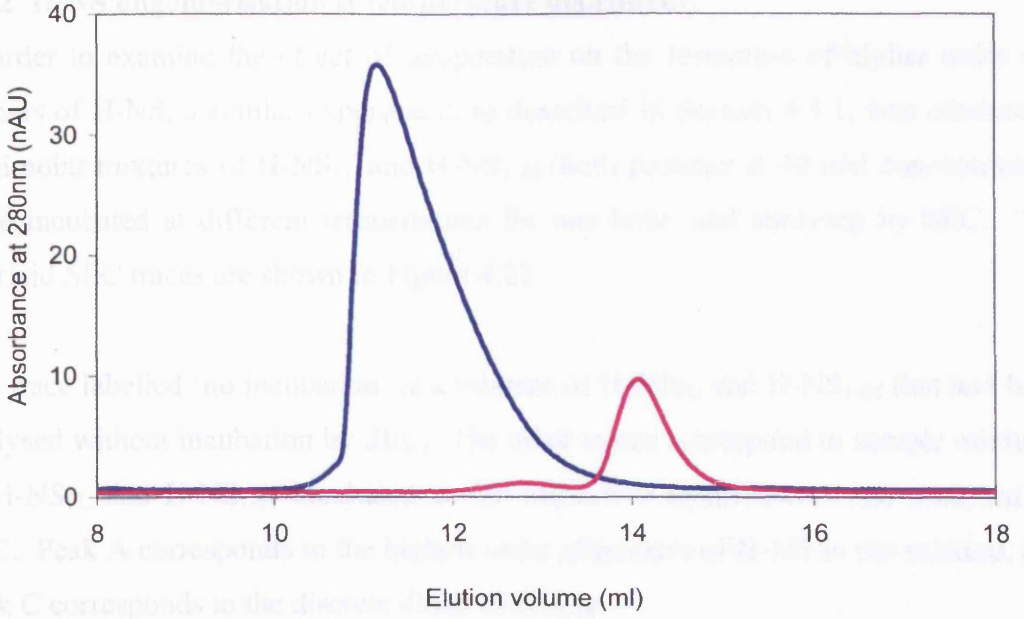


Figure 4.21
SEC traces of H-NS_{FL} (—) and H-$NS_{1\text{-}64}$ (—), with both proteins at 40 μM concentration. The protein samples were analysed by SEC independently, and the SEC traces overlaid. Differences in peak area reflect differences in per-molar extinction coefficients of the two proteins.

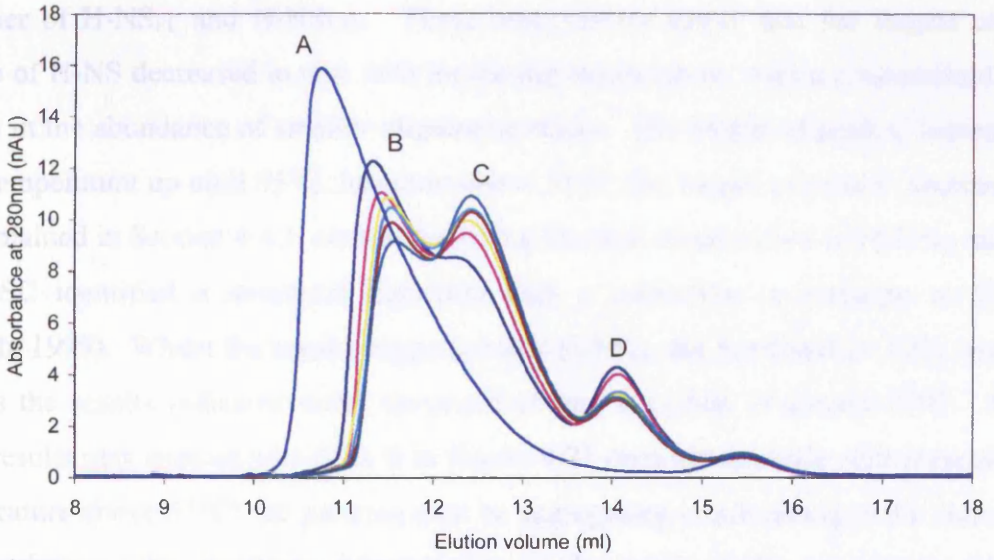


Figure 4.22
SEC traces of equimolar mixtures of H-$NS_{1\text{-}64}$ and H-NS_{FL} (at 40 μM concentration) incubated at 45°C for varying lengths of time (from 1 to 24 hours). The following traces are shown: mixtures of H-$NS_{1\text{-}64}$ and H-NS_{FL} co-incubated for 1 hour (—), 2 hours (—), 3 hours (—), 4 hours (—), 6 hours (—) and 24 hours (—). A trace corresponding to a sample containing only 40 μM H-NS_{FL} (—) is also shown.

4.5.2 H-NS oligomerisation is temperature dependent

In order to examine the effect of temperature on the formation of higher order oligomers of H-NS, a similar experiment, as described in Section 4.5.1, was conducted. Equimolar mixtures of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$ (both proteins at 40 μM concentration) were incubated at different temperatures for one hour, and analysed by SEC. The overlaid SEC traces are shown in Figure 4.23.

The trace labelled 'no incubation' is a mixture of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$ that had been analysed without incubation by SEC. The other traces correspond to sample mixtures of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$ incubated at the respective temperatures and analysed by SEC. Peak A corresponds to the highest order oligomers of H-NS in the mixture, and peak C corresponds to the discrete dimer $H\text{-}NS_{1\text{-}64}$.

$H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$ interacted in a temperature dependent manner, with the collection of peaks A (corresponding to the largest molecular weight complexes of H-NS) appearing at higher elution volumes with increasing temperature. As the incubation temperature is increased, peaks A and C also decreased in intensity: this was mirrored by an increase in the height of peak B, which is likely to correspond to the heterotetramer of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$. These observations imply that the largest complexes of H-NS decreased in size with increasing temperature, with a concomitant increase in the abundance of smaller oligomeric states. The height of peak C increases with temperature up until 55°C, however above 55°C, the height of peak C decreases. As explained in Section 4.4.1, studies involving thermal denaturation of $H\text{-}NS_{FL}$ using the DSC identified a structural transition with a maximum in enthalpy at 55°C (Smyth, 1999). Whilst the results suggested that $H\text{-}NS_{FL}$ did not denature fully, nonetheless the results indicated some structural change occurred at around 55°C. The DSC results may explain why peak B in Figure 4.23 does not increase with increasing temperature above 55°C: the proteins may be aggregating or adsorbing to the sides of the incubation tube, rendering interpretation of the result of the experiment above 55°C difficult. Nonetheless, between the ranges 35°C to 55°C, an apparent increase in the peak corresponding to the heterotetramer of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$ is observed in the SEC traces with increasing temperature.

4.5.3 H-NS_{FL} forms oligomers in a concentration dependent manner

In previous studies involving SEC, H-NS_{FL} has been shown to form higher order oligomers at least as large as icosamers in a concentration dependent fashion (Smyth *et al.*, 2000), hence in a solution of fixed H-NS_{FL} concentration, increasing concentrations of H-NS_{1-64} would compete for the oligomerisation sites found on H-NS_{FL}.

According to the 'head-to-tail' model (see Section 1.9.4), H-NS_{1-64} interacts and forms a complex with H-NS_{FL} via its 'tail' structure, residing between residues 1 to 64, with the 'head' (residues 65 to 89) of an adjacent H-NS_{FL} dimer. However, lacking the residues 65 to 89, H-NS_{1-64} is unable to interact further with other dimers of H-NS_{FL}. In effect, H-NS_{1-64} inhibits the formation of higher order oligomers. This can be seen in the following experiment where samples containing increasing ratios of H-NS_{1-64}: H-NS_{FL} are analysed by SEC. Mixtures containing H-NS_{FL} at 40 μM and H-NS_{1-64} at a range of concentrations between 40 μM and 1.92 mM were prepared and incubated for 15 hours at 4°C prior to analysis. The overlaid SEC traces are shown in Figure 4.24. The collection of peaks labelled A correspond to the largest oligomeric species of H-NS in the sample; peak B most likely corresponds to the heterotetramer of H-NS_{FL} dimer and H-NS_{1-64} dimer; and peak C corresponds to discrete dimeric H-NS_{1-64}.

As the ratio of H-NS_{1-64}: H-NS_{FL} is increased, the amount of the heterotetramer of H-NS_{FL} and H-NS_{1-64} (corresponding to peak B) increases. This is consistent with the proposal that H-NS_{1-64} directly competes for the oligomerisation sites present on H-NS_{FL}. However due to the inability of H-NS_{1-64} to promote further oligomerisation due to the lack of the second oligomerisation interface found between residues 65 and 89, H-NS_{1-64} effectively acts to inhibit the formation of higher order oligomers. This is shown by the small, but nonetheless consistent, increase of elution volume of the largest oligomeric species (corresponding to the collection of peaks labelled A); in other words, the higher order oligomeric states are decreasing in size as the ratio of H-NS_{1-64}: H-NS_{FL} is increased.

There is an apparent linear relationship between the increase in amount of the H-NS_{FL}/H-NS_{1-64} heterotetramer (corresponding to peak B) and the ratio of H-NS_{1-64}: H-NS_{FL} concentration. The interaction between H-NS_{FL} and H-NS_{1-64} has not reached

equilibrium under the conditions of the experiment. There is no apparent reason why the binding affinity of one dimer of H-NS_{1-64} with a dimer of H-NS_{FL} should significantly differ from the binding affinity of two H-NS_{FL} molecules interacting with each other. If the assumption is made that these two binding affinities are equal, it is highly unlikely that ratios of H-NS_{1-64}: H-NS_{FL} concentration as high as 48 are needed to achieve saturation of H-NS_{FL} oligomerisation sites with H-NS_{1-64}. Instead, it is more likely that either the incubation period of 15 hours was not long enough to allow equilibrium of the interaction of H-NS_{FL} and H-NS_{1-64} to occur, or that incubation temperature was too low. Indeed, in Figure 4.22, equilibrium of the interaction between H-NS_{FL} and H-NS_{1-64} is reached after approximately 5 hours at 45°C.

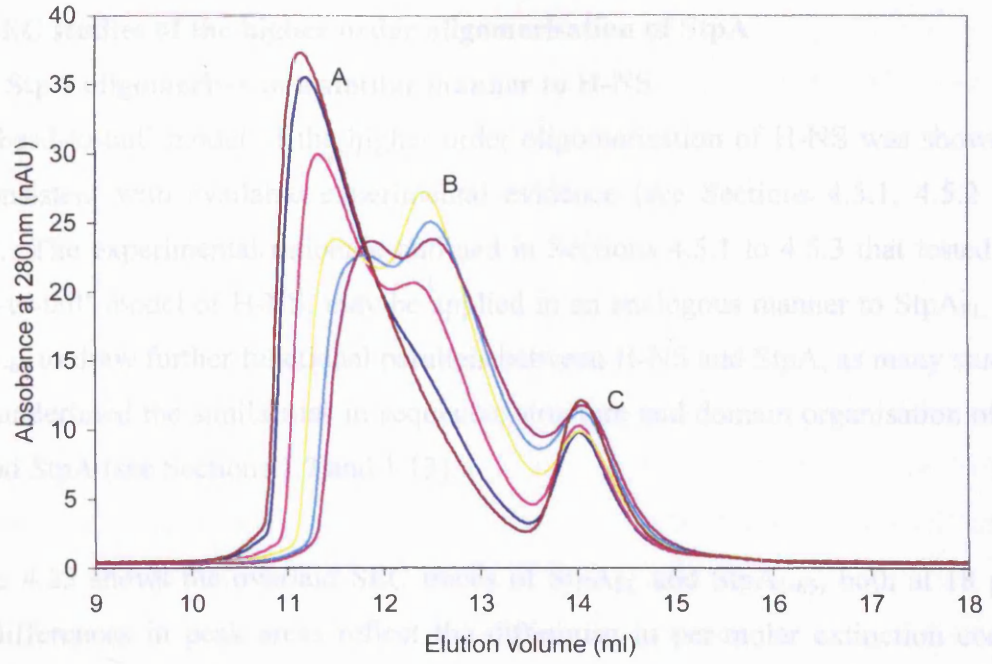


Figure 4.23
SEC traces of equimolar mixtures of H-NS_{1-64} and H-NS_{FL} (at 40 μM concentration) incubated for 1 hour at varying temperatures (from 35°C to 75°C). The following traces are shown: mixtures of H-NS_{1-64} and H-NS_{FL} incubated at 35°C (—), 45°C (—), 55°C (—), 65°C (—) and 75°C (—). A trace corresponding to an unincubated sample is also shown (—).

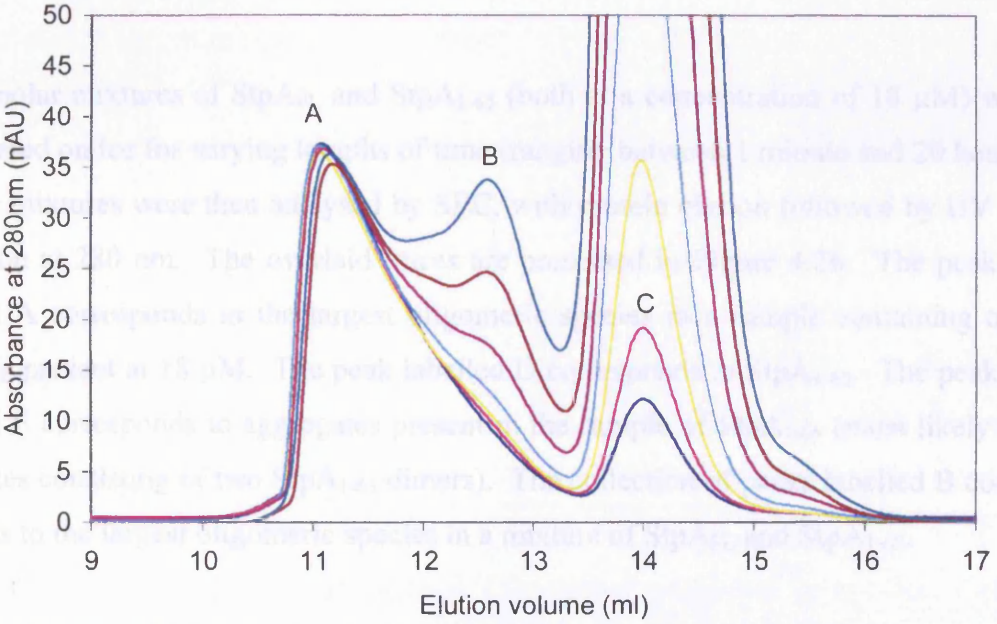


Figure 4.24
SEC traces of mixtures of H-NS_{1-64} and H-NS_{FL} at varying molar ratios incubated for 15 hours at 4°C). Traces from samples containing the following ratios of H-NS_{1-64}: H-NS_{FL} are shown: 1 (—), 2 (—), 4 (—), 8 (—), 16 (—), 32 (—) and 48 (—). Note that H-NS_{FL} concentration was set at 40 μM.

4.6 SEC studies of the higher order oligomerisation of StpA

4.6.1 StpA oligomerises in a similar manner to H-NS

The 'head-to-tail' model of the higher order oligomerisation of H-NS was shown to be consistent with available experimental evidence (see Sections 4.5.1, 4.5.2 and 4.5.3). The experimental rationale outlined in Sections 4.5.1 to 4.5.3 that tested the 'head-to-tail' model of H-NS, may be applied in an analogous manner to $StpA_{FL}$ and $StpA_{1-65}$ to draw further functional parallels between H-NS and StpA, as many studies have underlined the similarities in sequence, structure and domain organisation of H-NS and StpA (see Sections 1.7 and 1.13).

Figure 4.25 shows the overlaid SEC traces of $StpA_{FL}$ and $StpA_{1-65}$, both at 18 μM. The differences in peak areas reflect the difference in per-molar extinction coefficients of $StpA_{FL}$ and $StpA_{1-65}$. According to the rationale outlined in Section 4.5.1, a mixture of $StpA_{FL}$ and $StpA_{1-65}$ should reveal a chromatogram that is the summation of the two traces in Figure 4.25, if the two proteins do not interact with each other. On the other hand, interaction between $StpA_{FL}$ and $StpA_{1-65}$ should cause the chromatogram of a mixture of the two proteins to have a significantly different shape to a simple summation of the traces shown in 4.21.

Equimolar mixtures of $StpA_{FL}$ and $StpA_{1-65}$ (both at a concentration of 18 μM) were incubated on ice for varying lengths of time (ranging between 1 minute and 20 hours). These mixtures were then analysed by SEC, with protein elution followed by UV absorption at 280 nm. The overlaid traces are presented in Figure 4.26. The peak labelled A corresponds to the largest oligomeric species in a sample containing only $StpA_{FL}$ present at 18 μM. The peak labelled D corresponds to $StpA_{1-65}$. The peak labelled E corresponds to aggregates present in the sample of $StpA_{1-65}$ (most likely aggregates consisting of two $StpA_{1-65}$ dimers). The collection of peaks labelled B corresponds to the largest oligomeric species in a mixture of $StpA_{FL}$ and $StpA_{1-65}$.

The addition of $StpA_{1-65}$ to a sample of $StpA_{FL}$ causes the distribution of oligomeric species sizes to change rapidly. This is shown by an increase in elution volume from the largest oligomeric species in a sample of $StpA_{FL}$ (peak A) to those found in the mixtures of $StpA_{FL}$ and $StpA_{1-65}$ (the collection of peaks labelled B). The addition of $StpA_{1-65}$ to a sample of $StpA_{FL}$ has significantly decreased the sizes of the higher or-

der oligomeric species observed by SEC. All interactions between $StpA_{FL}$ and $StpA_{1-65}$ take place and reach a state of equilibrium rapidly. In other words, all detectable changes in the distribution of oligomeric species occur within one minute of addition of $StpA_{1-65}$ to $StpA_{FL}$. Incubation for a further 20 hours has no detectable effect on the distribution of oligomeric species.

By analogy to H-NS, $StpA_{FL}$ has both a 'head' (a structural determinant found somewhere between residues 66 and 90) and a 'tail' (a structural determinant found somewhere in the N-terminal domain, between residues 1 to 65). $StpA_{1-65}$ contains only the 'tail' region. By binding via its 'tail' region with the free 'head' region of another $StpA_{FL}$ dimer, $StpA_{1-65}$ can bind to $StpA_{FL}$ to form a heterotetrameric complex of a $StpA_{FL}$ dimer and a dimer of $StpA_{1-65}$. As the $StpA_{FL}$ dimer in this complex retains a free 'tail' region, this heterotetrameric complex is free to interact with other dimers of $StpA_{FL}$ and form higher order oligomers. However, as $StpA_{1-65}$ lacks the 'head' region (as it lacks the residues 66 to 90), it is unable to contribute to the formation of higher order oligomers. In effect, the binding of $StpA_{1-65}$ to an oligomer of $StpA_{FL}$ has 'capped' one end of the oligomer, preventing further units of $StpA_{FL}$ dimers from binding. This 'capping' effect of $StpA_{1-65}$ on an oligomer of $StpA_{FL}$ will reduce the overall size of oligomer seen in this mixture of $StpA_{FL}$ and $StpA_{1-65}$.

In similar experiments involving $H\text{-}NS_{FL}$ and $H\text{-}NS_{1-64}$ (see Sections 4.5.1, 4.5.2 and 4.5.3), the co-incubation of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1-64}$ resulted in a peak at an elution volume of 12.4 ml that most likely corresponded to a heterotetramer of a $H\text{-}NS_{FL}$ dimer and a $H\text{-}NS_{1-64}$ dimer. However, an analogous peak corresponding to a heterotetramer of $StpA_{FL}$ and $StpA_{1-65}$ is not observed in Figure 4.26. This implies a detectable amount of the heterotetrameric complex of $StpA_{FL}$ and $StpA_{1-65}$ is not produced. There is a very small peak at an elution volume of approximately 12.5 ml (labelled 'F'); however, this peak may correspond to aggregated $StpA_{1-65}$, which is also present in a sample of $StpA_{1-65}$ as a peak labelled 'E'. As the 'pull-down' assays described in section 4.4.2 suggest, monomers may exchange rapidly between the coiled-coils of StpA. This in turn may result in a more polydisperse population of heterooligomers of $StpA_{FL}$ and $StpA_{1-65}$ in the mixtures analysed by SEC, leading to the broad distribution of oligomeric species as shown in Figure 4.26. However, the resolution of the Superose 12 SEC column is insufficient to characterise the various oligomeric states.

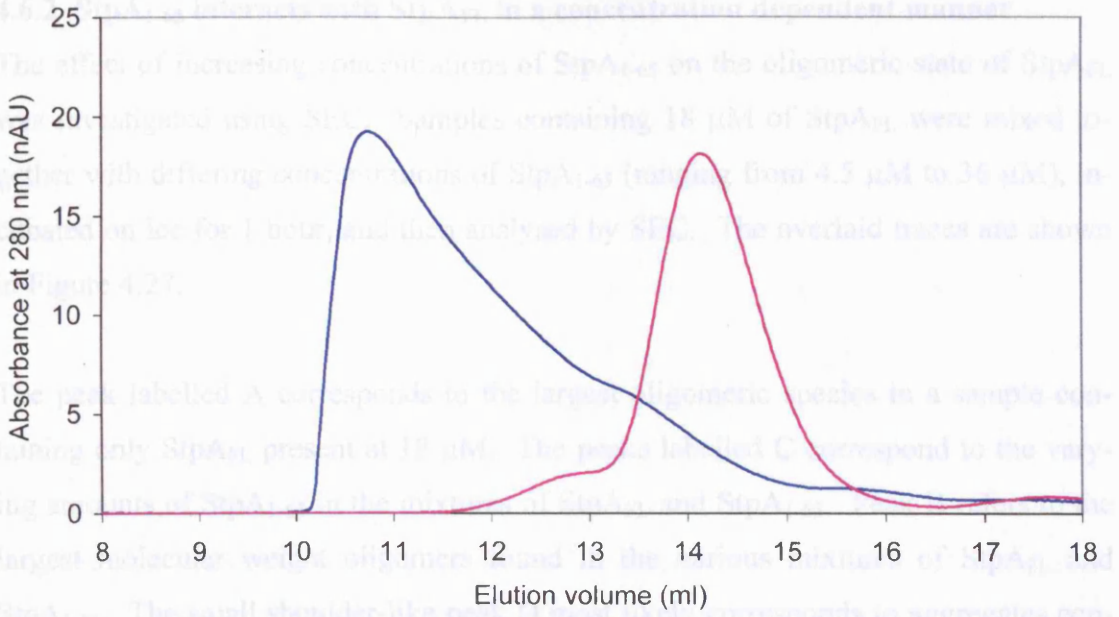


Figure 4.25
SEC traces of $StpA_{FL}$ (—) and $StpA_{1-65}$ (—), with both at 18 μM concentration. The protein samples were analysed by SEC independently, and the SEC traces overlaid. Differences in peak area reflect differences in per-molar extinction coefficients of the two proteins.

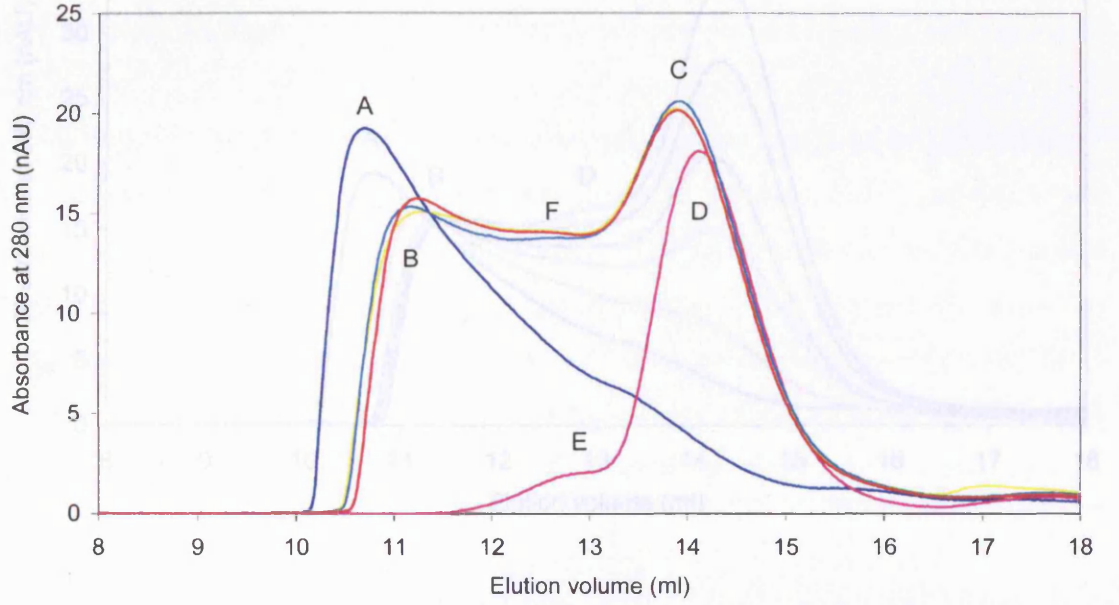


Figure 4.26
SEC traces of equimolar mixtures of $StpA_{1-65}$ and $StpA_{FL}$ (at 18 μM concentration) incubated on ice for varying lengths of time (from 1minute to 20 hours). The following traces are shown: mixtures of $StpA_{1-65}$ and $StpA_{FL}$ incubated for 1 minute (—), 30 minutes (—), and 20 hours (—). Traces corresponding to a sample containing only 18 μM $StpA_{1-65}$ (—) and a sample contending only 18 μM $StpA_{1-65}$ (—) are also shown.

4.6.2 $StpA_{1-65}$ interacts with $StpA_{FL}$ in a concentration dependent manner

The effect of increasing concentrations of $StpA_{1-65}$ on the oligomeric state of $StpA_{FL}$ was investigated using SEC. Samples containing 18 μM of $StpA_{FL}$ were mixed together with differing concentrations of $StpA_{1-65}$ (ranging from 4.5 μM to 36 μM), incubated on ice for 1 hour, and then analysed by SEC. The overlaid traces are shown in Figure 4.27.

The peak labelled A corresponds to the largest oligomeric species in a sample containing only $StpA_{FL}$ present at 18 μM. The peaks labelled C correspond to the varying amounts of $StpA_{1-65}$ in the mixtures of $StpA_{FL}$ and $StpA_{1-65}$. Peak B refers to the largest molecular weight oligomers found in the various mixtures of $StpA_{FL}$ and $StpA_{1-65}$. The small shoulder-like peak D most likely corresponds to aggregates consisting of two $StpA_{1-65}$ dimers, equivalent to peak E in Figure 4.26.

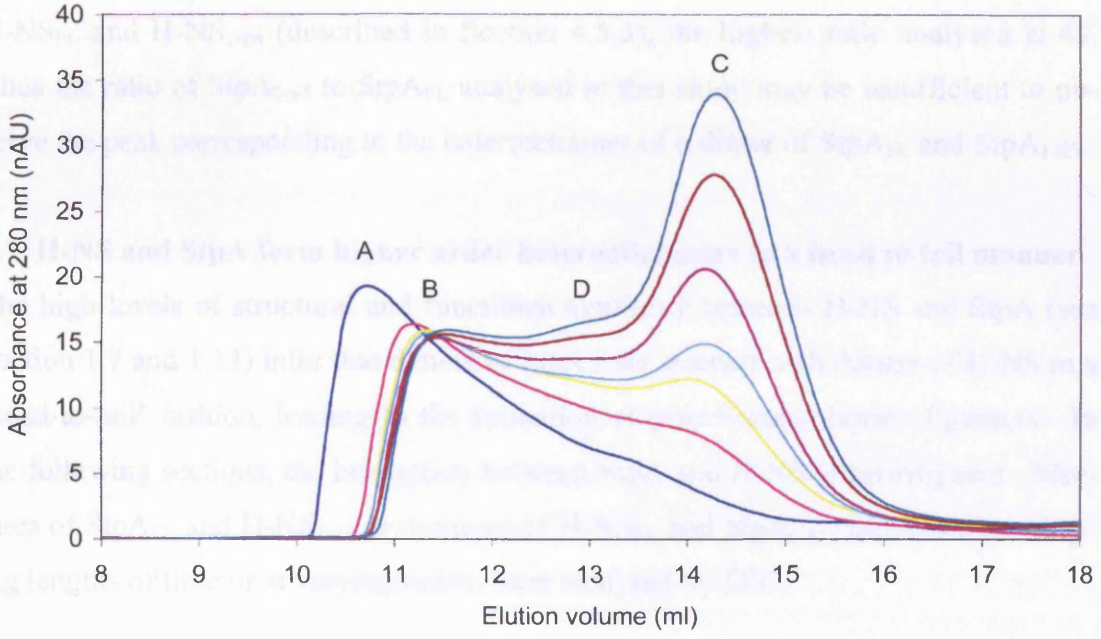


Figure 4.27
SEC traces of mixtures of $StpA_{1-65}$ and $StpA_{FL}$ at varying molar ratios incubated for 1 hour on ice. Traces from samples containing the following ratios of $StpA_{1-65}$: $StpA_{FL}$ are shown: 0.25 (—), 0.5 (—), 0.67 (—), 1 (—), 1.5 (—) and 2 (—). Note that $StpA_{FL}$ concentration was set at 18 μM (—).

Upon addition of $StpA_{1-65}$ to a sample of $StpA_{FL}$ there is a decrease in size of the largest molecular weight complex. This can be seen in the increase of the elution volume of the largest oligomeric complexes of the mixtures of $StpA_{FL}$ and $StpA_{1-65}$ (corresponding to the collection of peaks labelled B) compared to the largest molecular weight complexes of a sample of $StpA_{FL}$ (corresponding to peak A). A large decrease in the size of the largest molecular weight complex is seen upon the addition of a modest amount of $StpA_{1-65}$ (4.5 μM, or a quarter of the concentration of $StpA_{FL}$). Incremental increases of $StpA_{1-65}$ concentration (up to 36 μM) in a mixture with $StpA_{FL}$ yield smaller increases in elution volume of the largest oligomeric species.

As is the case with the experiment described in Section 4.6.1, there is no detectable peak corresponding to the heterotetramer of a dimer of $StpA_{FL}$ and $StpA_{1-65}$. This may be a consequence of the rapid exchange of the monomers of $StpA_{FL}$ and $StpA_{1-65}$, leading to a polydisperse population of heterooligomers that is insufficiently resolved on the Superose 12 SEC column. Also note that the highest ratio of N-terminal domain protein to full-length protein is 2, whereas with the equivalent experiment with $H\text{-}NS_{FL}$ and $H\text{-}NS_{1-64}$ (described in Section 4.5.3), the highest ratio analysed is 48. Thus the ratio of $StpA_{1-65}$ to $StpA_{FL}$ analysed in this study may be insufficient to observe the peak corresponding to the heterotetramer of a dimer of $StpA_{FL}$ and $StpA_{1-65}$.

4.7 H-NS and StpA form higher order heterooligomers in a head to tail manner

The high levels of structural and functional similarity between H-NS and StpA (see Section 1.7 and 1.13) infer that dimers of StpA may interact with dimers of H-NS in a 'head-to-tail' fashion, leading to the formation of polydisperse heterooligomers. In the following sections, the interaction between StpA and H-NS is investigated. Mixtures of $StpA_{FL}$ and $H\text{-}NS_{1-64}$ or mixtures of $H\text{-}NS_{FL}$ and $StpA_{1-65}$, incubated for varying lengths of time or at varying ratios, were analysed by SEC.

4.7.1 H-NS and StpA form heterooligomers rapidly in a head-to-tail manner

Equimolar mixtures of H-NS_{FL} and $StpA_{1-65}$ (both at 40 μM concentration) were incubated for varying lengths of time (from 1 minute to 6 hours) on ice and analysed by SEC. The overlaid SEC traces are shown in Figure 4.28. The SEC trace with the asymmetric peak labelled 'A' corresponds to a sample containing only H-NS_{FL}. The largest oligomeric species in a sample of H-NS_{FL} elute with a maximum at an elution volume of 11 ml. The SEC trace corresponding to a sample containing only $StpA_{1-65}$ is labelled 'C' with a maximum at an elution volume of 14 ml. The collection of peaks labelled 'B' corresponds to the heterotetramer consisting of a dimer of H-NS_{FL} and a dimer of $StpA_{1-65}$ in the various mixtures of $StpA_{1-65}$ and H-NS_{FL}.

The elution volume of the largest oligomeric species seen in a sample containing only H-NS_{FL} (corresponding to peak A) increases significantly on the addition of $StpA_{1-65}$ (corresponding to the collection of peaks labelled B). This suggests that oligomeric states in a mixture of H-NS_{FL} and $StpA_{1-65}$ are much smaller in the oligomers is a sample of H-NS_{FL} alone. Upon mixing, both peaks A and C decrease in size significantly: there is a concomitant increase in the peaks B, suggesting that much of the H-NS_{FL} and $StpA_{1-65}$ in the mixture elutes as a heterogeneous complex.

The experiment described above (with corresponding results shown in Figure 4.28) explored the interaction of $StpA_{1-65}$ with H-NS_{FL}, in other words, the interaction of the 'tail' region of $StpA_{1-65}$ (i.e. the N-terminal region of $StpA_{1-65}$) with the 'head' region of H-NS_{FL} (i.e. the residues 65 to 89). The following experiment investigates the reversed situation, where the interaction of the equivalent 'tail' region of H-NS_{1-64} with the 'head' region of $StpA_{FL}$ is analysed. Equimolar mixtures of $StpA_{FL}$ and H-NS_{1-64} (both at 18 μM concentration) were incubated on ice for varying lengths of time (from 1 minute to 18 hours) and then analysed by SEC. The overlaid SEC traces are shown in Figure 4.29.

The SEC trace of a sample containing only $StpA_{FL}$ exhibits a broad asymmetric peak (labelled 'A'), the maximum of which corresponds to the largest oligomeric complexes in the sample. The addition of H-NS_{1-64} to a sample of $StpA_{FL}$ causes the largest oligomeric species (corresponding to peak A) to drop dramatically in size. This is shown by an increase in the elution volume of the largest oligomeric species from

peak A to the small, badly resolved shoulder-peak B upon addition of H-NS_{1-64} to the sample of $StpA_{FL}$. A discrete peak, found at an elution volume of 12.4 ml (labelled peak C), is formed as a result of the interaction of H-NS_{1-64} and $StpA_{FL}$, and is likely to correspond to a heterotetramer of $StpA_{FL}$ and H-NS_{1-64}. The badly resolved shoulder-peak D corresponds to discrete H-NS_{1-64} dimer that has not formed a complex with $StpA_{FL}$.

Upon addition of $StpA_{1-65}$ into a solution of H-NS_{FL}, the sample equilibrates very rapidly: no significant change in the SEC traces is observed after 10 minutes of incubation on ice. The distribution of oligomeric species rapidly changes upon addition of $StpA_{1-65}$ to the sample of H-NS_{FL}, as shown in Figure 4.28. Similarly the addition of H-NS_{1-64} to a sample of $StpA_{FL}$ causes a drastic and immediate change in the distribution of oligomeric species, as shown in Figure 4.29. The redistribution of oligomeric species apparently reaches equilibrium within 1 minute after the mixing of $StpA_{FL}$ and H-NS_{1-64}. In comparison, the redistribution of the population of oligomeric species upon the addition of H-NS_{1-64} to a sample of H-NS_{FL} (as shown in Section 4.5.1, Figure 4.22) apparently reaches equilibrium only after approximately 5 hours.

The Figures 4.28 and 4.29 show that the oligomerisation processes between H-NS and StpA are very rapid compared to the equivalent processes leading to the formation of homomeric complexes of H-NS, as shown in Figure 4.22. This suggests that any free StpA dimer *in vivo* immediately forms heteromeric complexes with H-NS in the bacterial cell, and that free StpA is unlikely to be found in the cell. Although the transcription from the *stpA* gene is usually strongly repressed by H-NS (Zhang *et al.*, 1996; Free and Dorman, 1997), basal levels of transcription results in small amounts of StpA in the cell in any given situation. If indeed large polydisperse oligomers of H-NS exist *in vivo* as biophysical experiments suggest (Smyth *et al.*, 2000), it is highly likely that StpA will be found in the large oligomers of H-NS. This is in agreement with studies that suggest that in the absence of H-NS, StpA is rapidly degraded by Lon protease (Johansson and Uhlin, 1999; Johansson *et al.*, 2001). The rapid interaction of any newly synthesised StpA with H-NS in the bacterial cell would ensure prevention from degradation of StpA by Lon protease.

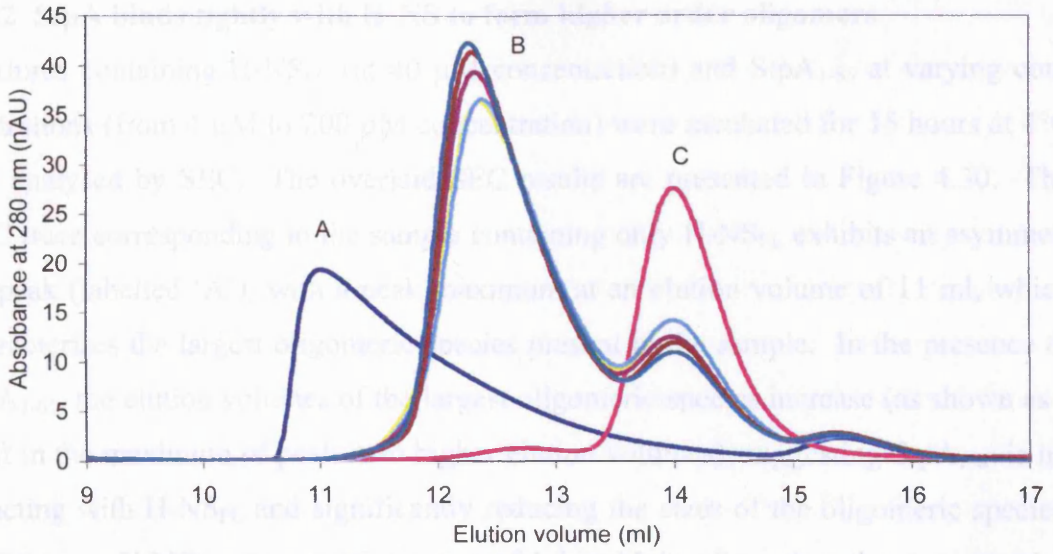


Figure 4.28
SEC traces of equimolar mixtures of $StpA_{1-65}$ and $H\text{-}NS_{FL}$ (at 40 μM concentration) incubated on ice for varying lengths of time (from 1minute to 6 hours). The following traces are shown: mixtures of $StpA_{1-65}$ and $H\text{-}NS_{FL}$ incubated for 1 minute (—), 5 minutes (—), 10 minutes (—), 1 hour (—), and 6 hours (—). Traces corresponding to a sample containing only 40 μM $StpA_{1-65}$ (—) and a sample contending only 40 μM $H\text{-}NS_{FL}$ (—) are also shown.

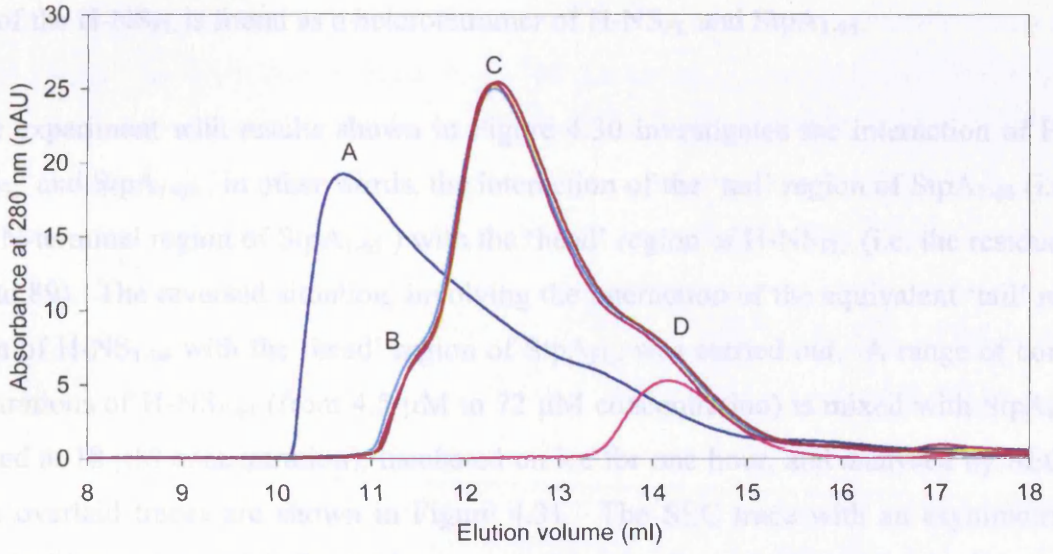


Figure 4.29
SEC traces of equimolar mixtures of $H\text{-}NS_{1-64}$ and $StpA_{FL}$ (both at 18 μM concentration) incubated on ice for varying lengths of time (from 1minute to 18 hours). The following traces are shown: mixtures of $H\text{-}NS_{1-64}$ and $StpA_{FL}$ incubated for 1 minute (—), 10 minutes (—), 30 minutes (—), and 18 hours (—). Traces corresponding to a sample containing only 18 μM $H\text{-}NS_{1-64}$ (—) and a sample contending only 18 μM $StpA_{FL}$ (—) are also shown.

4.7.2 StpA binds tightly with H-NS to form higher order oligomers

Mixtures containing H-NS_{FL} (at 40 μM concentration) and $StpA_{1-65}$ at varying concentrations (from 4 μM to 200 μM concentration) were incubated for 15 hours at 4°C and analysed by SEC. The overlaid SEC results are presented in Figure 4.30. The SEC trace corresponding to the sample containing only H-NS_{FL} exhibits an asymmetric peak (labelled 'A'), with a peak maximum at an elution volume of 11 ml, which characterises the largest oligomeric species present in the sample. In the presence of $StpA_{1-65}$, the elution volumes of the largest oligomeric species increase (as shown as a shift in the maximum of peak A to higher elution volumes), suggesting $StpA_{1-65}$ is interacting with H-NS_{FL} and significantly reducing the sizes of the oligomeric species. At $StpA_{1-65}$: H-NS_{FL} concentration ratios of 1:1 or higher (i.e. when the concentration of $StpA_{1-65}$ is equal or higher to the concentration of H-NS_{FL}), peak A disappears almost completely (or is too small to be deconvolved from peak B). This decrease in size of higher order oligomers is matched by a concomitant increase in height of the discrete peak B eluting at 12.3 ml elution volume. The discrete peak B corresponds to the heterotetramer of a dimer of H-NS_{FL} and a dimer of $StpA_{1-65}$. The amount of the heterotetramer increases as the concentration of $StpA_{1-65}$ relative to the H-NS_{FL} is increased. When $StpA_{1-65}$ is present in equal or greater amounts than H-NS_{FL}, almost all of the H-NS_{FL} is found as a heterotetramer of H-NS_{FL} and $StpA_{1-65}$.

The experiment with results shown in Figure 4.30 investigates the interaction of H-NS_{FL} and $StpA_{1-65}$, in other words, the interaction of the 'tail' region of $StpA_{1-65}$ (i.e. the N-terminal region of $StpA_{1-65}$) with the 'head' region of H-NS_{FL} (i.e. the residues 65 to 89). The reversed situation, involving the interaction of the equivalent 'tail' region of H-NS_{1-64} with the 'head' region of $StpA_{FL}$, was carried out. A range of concentrations of H-NS_{1-64} (from 4.5 μM to 72 μM concentration) is mixed with $StpA_{FL}$ (fixed at 18 μM concentration), incubated on ice for one hour, and analysed by SEC. The overlaid traces are shown in Figure 4.31. The SEC trace with an asymmetric peak with maximum labelled 'A' corresponds to the largest oligomeric species in a sample containing only 18 μM $StpA_{FL}$. The discrete peaks labelled 'C' correspond to the varying amounts of H-NS_{1-64} in the mixtures of $StpA_{FL}$ and H-NS_{1-64}. In the presence of H-NS_{1-64}, the elution volume of the largest molecular weight complex increases (seen as a shift in peak A to higher elution volumes), suggesting H-NS_{1-64} interacts with $StpA_{FL}$ and reduces the sizes of the largest oligomeric species. The size

of the peak 'A' also decreases progressively as the concentration of H-NS_{1-64} is increased, eventually ending up as a small shoulder-peak (labelled 'D') of the much larger peak B. This decrease of peak height and shift to higher elution volumes of peak A is mirrored by an increase in the discrete peak B with an elution volume of 12.4 ml. This discrete peak B corresponds to the heterotetramer consisting of a dimer of $StpA_{FL}$ and a dimer of H-NS_{1-64}. At molar ratios of H-NS_{1-64}: $StpA_{FL}$ equal to or greater than 1, there are no significant differences in the shoulder-peak D or peak B, suggesting that further addition of H-NS_{1-64} results in little or no difference in the distribution of oligomeric species in a sample of $StpA_{FL}$ and H-NS_{1-64}. In other words, the binding sites on $StpA_{FL}$ equivalent to the 'head' (i.e. residues 66 to 90) are almost saturated with H-NS_{1-64} when the amount of H-NS_{1-64} in the sample is equal to or greater than $StpA_{FL}$, and that the predominant oligomeric species present in the sample is the heterotetramer of $StpA_{FL}$ and H-NS_{1-64} under the conditions specified.

Even the addition of small amounts of $StpA_{1-65}$ into a sample of H-NS_{FL} causes an abrupt decrease in the size of the largest oligomeric species, as shown in Figure 4.30. This decrease in the size of the large oligomers in the sample is mirrored by an increase in the peak corresponding to the heterotetramer of H-NS_{FL} and $StpA_{1-65}$. The majority of these changes in the distribution of the oligomers occur at molar ratios of $StpA_{1-65}$: H-NS_{FL} less than or equal to 1. In the reversed situation, where varying amounts of H-NS_{1-64} are incubated with $StpA_{FL}$ (as shown in Figure 4.31), a similar effect is observed, whereby the majority of the change in distribution in the oligomeric species occurs at molar ratios of H-NS_{1-64}: $StpA_{FL}$ less than or equal to 1. These results suggest that the vast majority of the N-terminal domain protein (either $StpA_{1-65}$ or H-NS_{1-64}) is bound to most available oligomerisation interfaces on the full-length protein (H-NS_{FL} or $StpA_{FL}$ respectively). The oligomerisation interfaces presented on the full-length protein (i.e. the 'head' regions of H-NS_{FL}, between residues 65-89 or of $StpA_{FL}$ between residues 66 to 90) are almost saturated when there are equal molar amounts of the N-terminal domain protein and the full-length protein. This suggests that the interaction between dimers of H-NS and StpA is stronger (or at least as strong) than the interaction of either two dimers of H-NS with each other, or two dimers of StpA with each other. As discussed at the end of Section 4.7.1, StpA dimer forms heteromeric complexes with H-NS *in vivo* immediately and with high affinity, and that homomeric complexes of StpA may not be found in the cell.

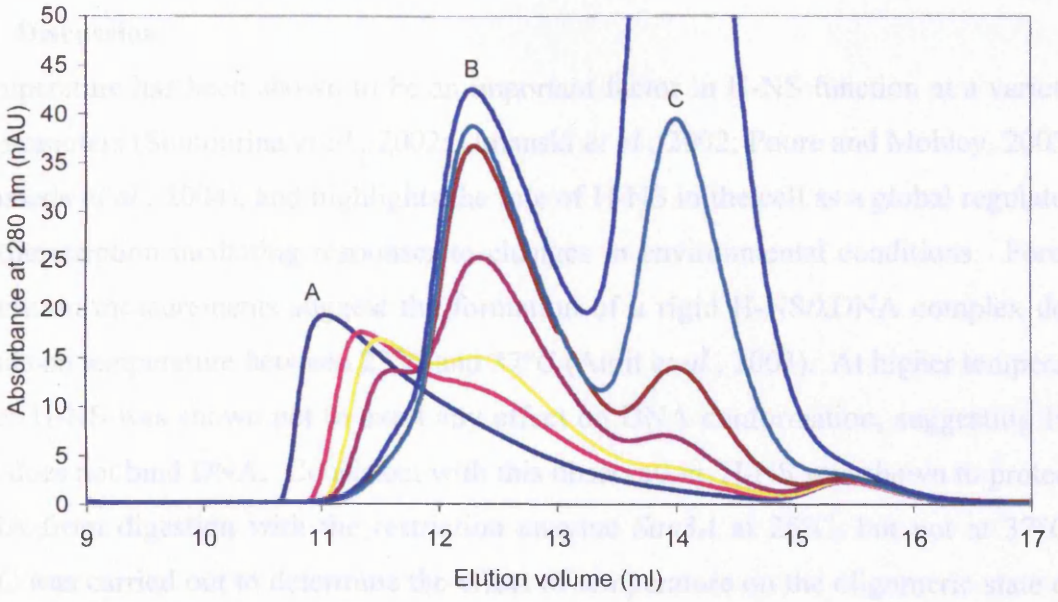


Figure 4.30
SEC traces of mixtures of H-NS_{FL} and $StpA_{1-65}$ at varying molar ratios incubated for 15 hours on ice. Traces from samples containing the following ratios of $StpA_{1-65}$: H-NS_{FL} are shown: 0.1 (or 4 μM $StpA_{1-65}$, —), 0.2 (or 8μM $StpA_{1-65}$, —), 0.5 (or 20 μM $StpA_{1-65}$, —), 1 (or 40 μM $StpA_{1-65}$, —), 2 (or 80 μM $StpA_{1-65}$, —) and 5 (or 200 μM $StpA_{1-65}$, —). A SEC trace of a sample containing only 40 μM H-NS_{FL} is shown (—).

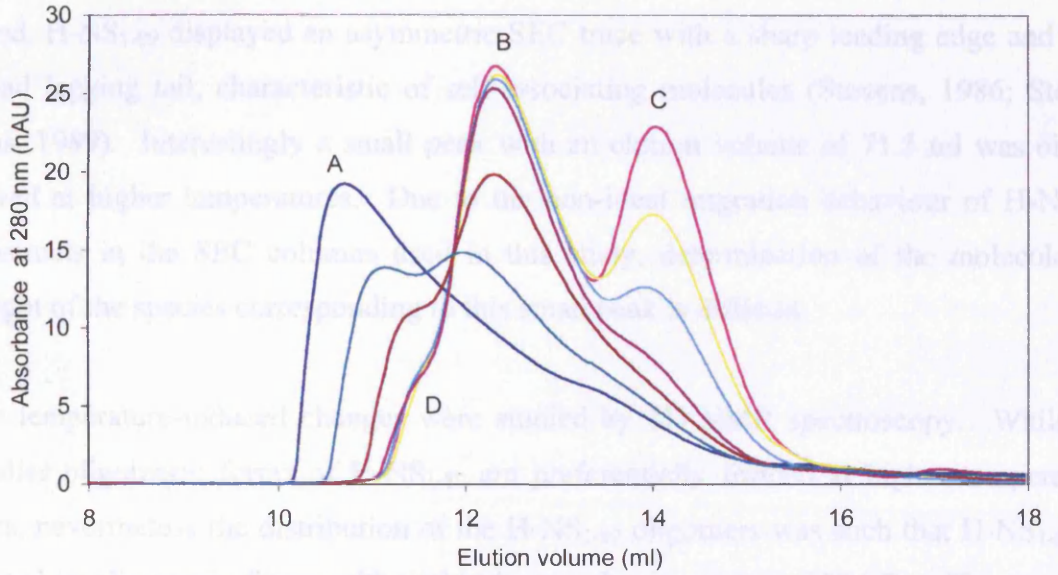


Figure 4.31
SEC traces of mixtures of $StpA_{FL}$ and H-NS_{1-64} at varying molar ratios incubated for 15 hours on ice. Traces from samples containing the following ratios of H-NS_{1-64}: $StpA_{FL}$ are shown: 0.25 (or 4.5 μM H-NS_{1-64}, —), 0.33 (or 6μM H-NS_{1-64}, —), 0.5 (or 9μM H-NS_{1-64}, —), 1 (or 18 μM H-NS_{1-64}, —), 2 (or 36 μM H-NS_{1-64}, —) and 4 (or 72 μM H-NS_{1-64}, —). A SEC trace of a sample containing only 18 μM $StpA_{FL}$ is also shown (—).

4.8 Discussion

Temperature has been shown to be an important factor in H-NS function at a variety of promoters (Soutourina *et al.*, 2002; Umanski *et al.*, 2002; Poore and Mobley, 2003; Prosseda *et al.*, 2004), and highlights the role of H-NS in the cell as a global regulator of transcription mediating responses to changes in environmental conditions. Force extension measurements suggest the formation of a rigid H-NS/λDNA complex depends on temperature between 23°C and 37°C (Amit *et al.*, 2003). At higher temperatures H-NS was shown not to exert any effect on DNA conformation, suggesting H-NS does not bind DNA. Consistent with this observation, H-NS was shown to protect DNA from digestion with the restriction enzyme *Sau3A* at 25°C, but not at 37°C. SEC was carried out to determine the effect of temperature on the oligomeric state of $H\text{-}NS_{1\text{-}89}$. The distribution of oligomeric species was clearly shown to depend on temperature, with higher temperatures favouring smaller oligomeric species. This is in direct disagreement with the findings of Ceschini *et al*, who found that higher temperatures favour the formation of larger oligomeric states of H-NS in studies involving large-zone gel-permeation chromatography (Ceschini *et al.*, 2000). Significant changes in the elution volume corresponding to the largest oligomeric species were observed in the temperature range investigated. Nonetheless, at all temperatures analysed, $H\text{-}NS_{1\text{-}89}$ displayed an asymmetric SEC trace with a sharp leading edge and a broad lagging tail, characteristic of self-associating molecules (Stevens, 1986; Stevens, 1989). Interestingly a small peak with an elution volume of 71.5 ml was observed at higher temperatures. Due to the non-ideal migration behaviour of H-NS constructs in the SEC columns used in this study, determination of the molecular weight of the species corresponding to this small peak is difficult.

The temperature-induced changes were studied by 1D NMR spectroscopy. Whilst smaller oligomeric forms of $H\text{-}NS_{1\text{-}89}$ are preferentially formed at higher temperatures, nevertheless the distribution of the $H\text{-}NS_{1\text{-}89}$ oligomers was such that $H\text{-}NS_{1\text{-}89}$ existed in oligomeric forms with molecular weights in excess of 30 kDa. The small peak with an elution volume of 71.5 ml corresponds to a species that is below the 30 kDa molecular weight limit imposed by the NMR technique. However, the concentration of this species may be below the detection limit of the NMR spectrometer. Otherwise, no differences whatsoever were observed between the 1D NMR spectra acquired at different temperatures.

A series of biophysical experiments were conducted to assess the structural and functional similarity of StpA with H-NS. The results of these experiments confirm the findings of previous studies which suggest that not only does StpA have high sequence, structure and domain organisation similarity to H-NS, but also that StpA is intimately linked with H-NS in mechanism of function.

The C-terminal domain of StpA was shown to exhibit radically different properties in solution compared to the N-terminal domain. The N-terminal domain of full-length StpA was effectively invisible by NMR spectroscopy, whilst the C-terminal domain of full-length StpA exhibited characteristics similar to the C-terminal domain of StpA in isolation (Figure 4.4). This suggests that StpA forms higher order oligomers, mediated by the N-terminal domain, whilst the C-terminal domain, linked by a long flexible linker, exhibits a higher degree of mobility relative to the N-terminal domain. In a analogous manner the C-terminal domain of H-NS exhibits different properties in solution compared to the N-terminal domain (Smyth *et al.*, 2000). CD experiments (Figure 4.6) and software used to predict propensity for coiled-coil formation (Figure 4.7) suggests the N-terminal domain of StpA forms an α-helix-rich coiled-coil domain, thereby providing further evidence of the high levels of structural homology between H-NS and StpA.

Experiments monitoring the thermally-induced denaturation by CD suggest that the N-terminal domain of StpA has a significantly higher T_m (midpoint of transition) than H-NS. Analysis of the change of the molar ellipticity at 220 nm, θ_{220nm}, of $StpA_{1-65}$ with respect to time suggested a cooperative, two-state unfolding process with a T_m of approximately 69°C. A previous CD study of H-NS indicated a T_m of 40.81°C (Schroder *et al.*, 2001). However, this CD study was based not on the N-terminal domain but on the full-length H-NS protein, thus thermal denaturation of either of the flexible linker region (which may exhibit a degree of secondary structure) or the C-terminal domain of H-NS may account for the lower T_m value determined for H-NS in this study. In addition, as a temperature range from only 24°C to 64°C was employed in this CD study, a denaturation event with a T_m greater than 64°C may have potentially been overlooked. Interestingly, experiments monitoring the thermal denaturation of the $StpA_{1-65}$ by CD showed that in low salt conditions (<50 mM NaCl), a dramatic change in secondary structure was observed above 75°C, where the α-helix rich

structure changed into one that was rich in β-sheet. This effect was pronounced when $StpA_{1-65}$ was present in 10 mM NaCl. This change in secondary structure was apparently irreversible: lowering of the temperature to below 75°C did not reverse the change in secondary structure. Under similar low salt conditions, an irreversible thermally-induced change in secondary structure of $H\text{-}NS_{FL}$ was not reported (Smyth, 1999; Schroder *et al.*, 2001). Both the T_m of $StpA_{1-65}$ and the irreversible thermally-induced change in secondary structure occur at non-physiological temperatures: therefore the relevance of the thermally-induced structural differences between H-NS and StpA determined by CD is not clear in the context of the structure-function relationship between H-NS and StpA.

In studies involving dominant-negative mutants of H-NS or StpA, a subset of mutants thought to be deficient in the formation of higher order oligomers were identified. This group of mutants include $H\text{-}NS_{1-64}$, $H\text{-}NS_{1-91}$ (containing only the first 91 residues of H-NS) and $H\text{-}NS_{E53G\ T55P}$ (where residues at the positions 53 and 55 were mutated to glycine and proline, respectively) (Williams *et al.*, 1996; Ueguchi *et al.*, 1996; Ueguchi *et al.*, 1997). Whilst competent in dimer formation via association of the N-terminal coiled-coil domain, these mutants were thought to exert their dominant-negative effect over wild-type H-NS (or StpA) by disrupting higher order oligomerisation. Examples of such dominant-negative mutants include $H\text{-}NS_{1-63}$ (containing only the first 63 N-terminal residues of H-NS) and $H\text{-}NS_{20-137}$ (missing the first 19 N-terminal amino acids) (Ueguchi *et al.*, 1997). There are two feasible mechanisms by which these H-NS mutants can exert their dominant-negative effect over wild-type H-NS (or StpA). According to the 'head-to-tail' model (Esposito *et al.*, 2002; Badaut *et al.*, 2002), oligomerisation-deficient mutants lacking the 'head' or 'tail' prevent the end-on-end oligomerisation of wild-type H-NS and limit the size of the oligomer formed, thereby resulting in the dominant-negative phenotype, as shown in Figure 4.32. However this explanation may be too simplistic: these dominant-negative mutants (e.g. $H\text{-}NS_{1-63}$ and $H\text{-}NS_{20-137}$) may interact with wild-type H-NS via the N-terminal coiled-coil region, forming heterodimers. The ensuing heterodimers of wild-type H-NS/ $H\text{-}NS_{1-63}$ and wild-type H-NS/ $H\text{-}NS_{20-137}$ would lack the functional 'head' and 'tail', respectively, again affecting oligomerisation of wild-type H-NS and resulting in the dominant-negative phenotype. The His-tag 'pull-down' assays investigated the degree to which the monomers of dimeric $H\text{-}NS_{1-64}$ can exchange, thereby

testing whether heterodimers (such as wild-type H-NS/ $H\text{-}NS_{1\text{-}64}$) may form. The results (shown in Figures 4.15 and 4.16) suggest that the N-terminal domain of H-NS presents a stable dimeric structure: only a small amount of exchange is observed at 4°C even after 40 hours. At 37°C, a significant level of exchange of the monomers of $H\text{-}NS_{1\text{-}64}$ is observed after 15 hours. An analogous His-tag 'pull-down' experiment was carried out on the N-terminal domain of StpA. Unfortunately, due to the high levels of contaminants in the samples of $StpA_{1\text{-}65}$ used, the results must be interpreted with care. Nonetheless, the results (shown in Figures 4.17 and 4.18) suggest that the monomers of the $StpA_{1\text{-}65}$ dimer exchange more rapidly than the monomers of $H\text{-}NS_{1\text{-}64}$: at 4°C and 37°C, complete exchange of the monomers of $StpA_{1\text{-}65}$ has apparently occurred in less than one hour. These results suggest that the dominant-negative effect displayed by H-NS (or StpA) defective in higher order oligomerisation can be mediated by the interaction of either the homodimer of the mutant or the heterodimer of wild-type H-NS/mutant H-NS with wild-type H-NS dimer, as shown in Figure 4.32.

Whilst previous studies have provided strong lines of evidence that suggest H-NS and StpA interact to form heterooligomers, it was not clear whether H-NS and StpA dimers interact via the 'head-to-tail' mechanism (in other words, H-NS dimers and StpA dimers interacting to form higher order heterooligomers) or H-NS and StpA monomers interact via the coiled-coil region to form a heterodimer. The His-tag 'pull-down' assays were modified to examine the interaction of H-NS and StpA N-terminal domains. Whilst the results (shown in Figures 4.19 and 4.20) suffer from the high levels of contamination present in samples of $StpA_{1\text{-}65}$, they suggest that heterodimers of $H\text{-}NS_{1\text{-}64}$ and $StpA_{1\text{-}65}$ form rapidly, i.e. at 4°C and 37°C, samples containing $StpA_{1\text{-}65}$ and $H\text{-}NS_{1\text{-}64}$ equilibrate in less than one hour.

The 'head-to-tail' model of H-NS oligomerisation is the most comprehensive model to date, and is consistent with all of the results presented in this chapter. $H\text{-}NS_{1\text{-}64}$ has previously been identified as a dominant-negative mutant (Williams *et al.*, 1996; Ueguchi *et al.*, 1996). The use of SEC to probe the interaction between $H\text{-}NS_{1\text{-}64}$ and $H\text{-}NS_{FL}$ provides insights into the mechanism of H-NS function and the dominant-negative effect. According to the 'head to tail' model, the interaction between $H\text{-}NS_{FL}$ and $H\text{-}NS_{1\text{-}64}$ involves the dissociation of two $H\text{-}NS_{FL}$ dimers involved in an

oligomer and the subsequent association of H-NS_{1-64} with one of the free interfaces on H-NS_{FL}. The apparent rate of this interaction was shown to increase with temperature between 35°C and 55°C (Figure 4.23). Above 55°C, factors such as denaturation of protein were suggested to interfere with the interaction. The interaction between H-NS_{FL} and H-NS_{1-64} was shown to increase as the concentration of H-NS_{1-64} was increased (shown in Figure 4.24). This is consistent with a previous study that showed H-NS_{FL} displays concentration-dependent oligomerisation (Smyth *et al.*, 2000). Interestingly, the apparent rate of interaction was shown to be very slow (Figure 4.22): samples of H-NS_{FL} and H-NS_{1-64} did not reach equilibrium for about 5 hours under the conditions tested. The interaction investigated by SEC between H-NS_{FL} and H-NS_{1-64} in Figures 4.22, 4.23 and 4.24 is characterised by the presence of a discrete peak at an elution volume of around 12.5 ml. An increase in this discrete peak is mirrored by the decrease in size of the largest oligomeric species. According to the 'head-to-tail' model, this peak corresponds to the heterotetramer consisting of a dimer of H-NS_{FL} and H-NS_{1-64}, which represents the smallest oligomeric species that can be formed between H-NS_{FL} and H-NS_{1-64}. Overexpression of H-NS_{1-64} in the bacterial cell may result in the majority of wild-type H-NS being found as a heterotetramer with H-NS_{1-64}, disrupting the wild-type function of H-NS and thus resulting in the dominant-negative effect of H-NS_{1-64}. However, the possibility of the formation of the heterodimer (H-NS_{FL}/H-NS_{1-64}) cannot be discounted, as shown in the His-tag 'pull-down' assays described in Section 4.4.1. If this is the case, then the discrete peak at 12.5 ml elution volume in the SEC traces (Figures 4.22, 4.23 and 4.24) corresponds to a heterotetramer of an H-NS_{FL} dimer and a heterodimer of H-NS_{FL}/H-NS_{1-64}). There is insufficient evidence to discriminate between the two models. In reality, the discrete peak at an elution volume of 12.5 ml is likely to correspond to a mixture of the heterotetramer consisting of H-NS_{FL} dimer and H-NS_{1-64} dimer and the heterotetramer consisting of H-NS_{FL} dimer and H-NS_{FL}/H-NS_{1-64} heterodimer.

The 'head-to-tail' model of H-NS was shown to adequately describe the higher order oligomerisation of StpA. In an analogous manner to H-NS, the addition of $StpA_{1-65}$ to a sample of $StpA_{FL}$ was shown to result in a decrease of the largest oligomeric species present in the sample, suggesting $StpA_{1-65}$ was interfering with $StpA_{FL}$ oligomerisation and limiting the size of the oligomers. The results suggest that the oligomerisation processes of $StpA_{FL}$ (in other words, dissociation and reassociation of StpA

dimers in the higher order oligomers) occur on faster time scales compared to H-NS, with sample mixture of $StpA_{FL}$ and $StpA_{1-65}$ equilibrating within minutes at temperatures as low as 0°C. Dissimilar to the findings of the SEC experiments of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1-64}$, no discrete peak at an elution volume of about 12.5 ml were observed in SEC traces of samples of $StpA_{FL}$ and $StpA_{1-65}$. This may result from the highly polydisperse distribution of oligomeric species in a sample of $StpA_{FL}$ and $StpA_{1-65}$, the separation of which is well beyond the resolution limit of the SEC column used.

The SEC experiments involving samples of either $StpA_{FL}$ and $H\text{-}NS_{1-64}$ or $H\text{-}NS_{FL}$ and $StpA_{1-65}$ suggest that H-NS and StpA can interact to form large polydisperse higher order heterooligomers. The affinity of an H-NS dimer for StpA dimer was shown to be higher (or at least equal) than the affinity of an H-NS dimer with another H-NS dimer. The apparent rate of the interaction between StpA and H-NS was shown to be fast, as sample mixtures of the relevant StpA and H-NS proteins equilibrated within minutes under the specified conditions. By comparison, a mixture of $H\text{-}NS_{FL}$ and $H\text{-}NS_{1-64}$ equilibrated after approximately 5 hours at higher temperatures. Therefore any free StpA dimer in solution is apparently rapidly integrated into the higher order oligomers of H-NS, leading to the formation of large polydisperse oligomers of H-NS and StpA. This is consistent with a study that showed StpA was rapidly degraded by Lon protease in an *hns* mutant (Johansson and Uhlin, 1999). The SEC experiments suggest that any expressed StpA is immediately incorporated into the higher order oligomers of H-NS *in vivo*, and that higher order oligomers consisting only of StpA are implausible.

The experiments described in this chapter highlight the intimate relationship between H-NS and StpA. In addition to the high levels of sequence and structural homology displayed by H-NS and StpA, the functional oligomerisation of the two proteins is inextricably linked. Higher order oligomers consisting of either pure $StpA_{FL}$ or pure $H\text{-}NS_{FL}$ are unlikely to exist *in vivo*. The unlikelihood of finding pure StpA in the cell may, in part, explain why no specific role for StpA has yet been found. The recent discovery of a third H-NS-like protein Sfh in *Shigella flexneri* strain 2457T suggests that the *in vivo* function of the H-NS-like higher order heterooligomers can be further modified by members of the H-NS-like family of proteins (Deighan *et al.*, 2003; Beloin *et al.*, 2003).

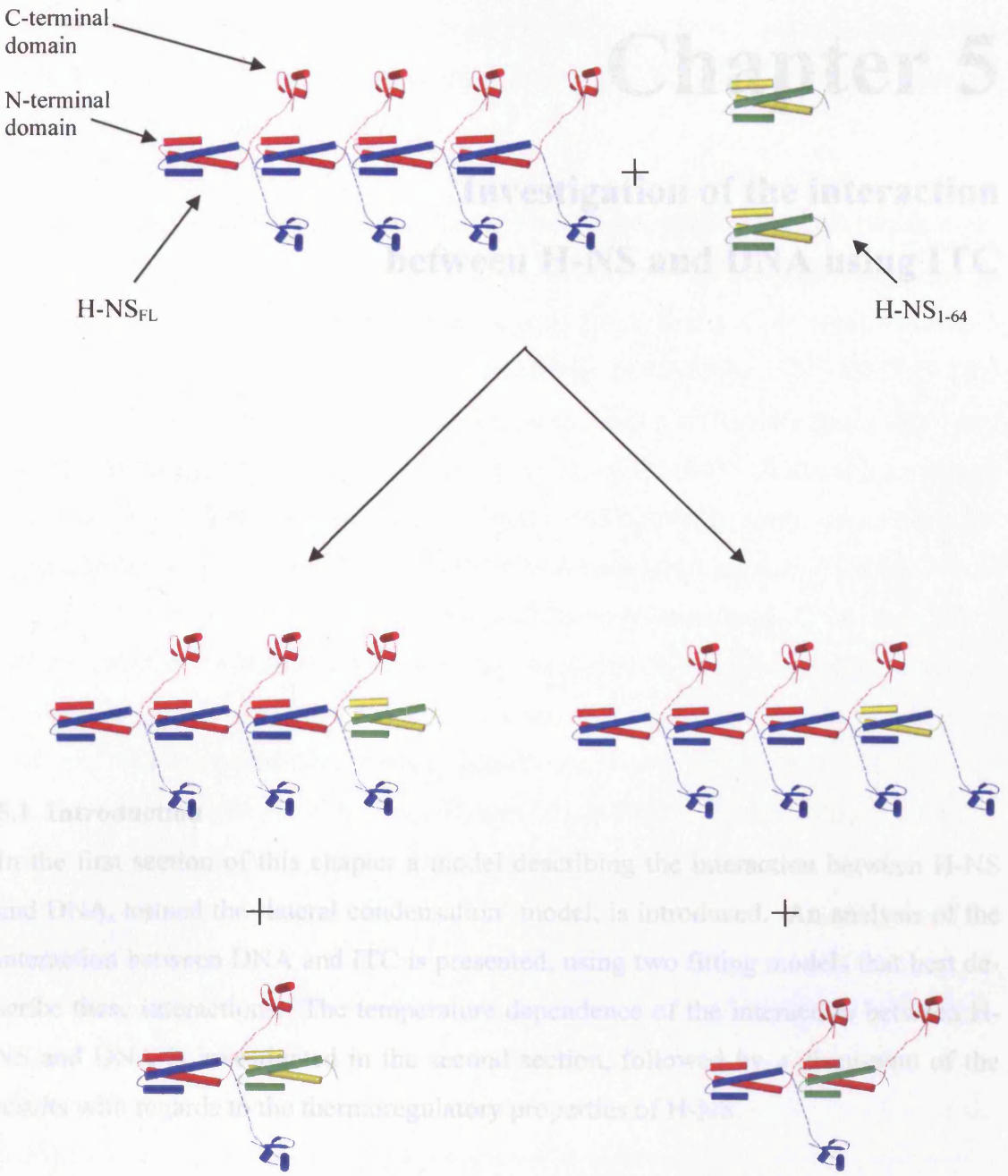


Figure 4.32
Schematic models describing the interaction of H-NS$_{FL}$ and H-NS$_{1-64}$. The monomers of H-NS$_{FL}$ are coloured red and blue, whilst the monomers of H-NS$_{1-64}$ are coloured green and yellow. The N-terminal domains of H-NS$_{FL}$ interact end-on-end to form higher order oligomers. H-NS$_{1-64}$can disrupt the formation of these oligomers by either of two mechanisms, shown schematically above. On the left hand side, H-NS$_{1-64}$ is shown interacting with H-NS$_{FL}$, resulting in the formation of various oligomers, one of which consists of a dimer of H-NS$_{FL}$ and a dimer of H-NS$_{1-64}$. On the right hand side, H-NS$_{1-64}$ is shown interacting with H-NS$_{FL}$ via the coiled-coil interface, resulting in the H-NS$_{FL}$/H-NS$_{1-64}$ heterodimer. This heterodimer then interacts with and disrupts the higher order oligomers of H-NS$_{FL}$.

Chapter 5

Investigation of the interaction between H-NS and DNA using ITC

5.1 Introduction

In the first section of this chapter a model describing the interaction between H-NS and DNA, termed the 'lateral condensation' model, is introduced. An analysis of the interaction between DNA and ITC is presented, using two fitting models that best describe these interactions. The temperature dependence of the interaction between H-NS and DNA is investigated in the second section, followed by a discussion of the results with regards to the thermoregulatory properties of H-NS.

Although efforts to characterise the DNA-binding properties of H-NS has been hampered by the inherent complexity of the mechanism involved, recent studies have provided valuable insights. The H-NS/DNA binding interaction is characterised by two key phenomena, the oligomerisation of H-NS, and the lateral condensation of two stretches of DNA (irrespective of sequence), shown schematically in Figure 5.1. This model is based on the 'head-to-tail' model of the oligomerisation of H-NS (see Section 1.9.4 and Chapter 4), where H-NS forms a 'protein scaffold' by the formation higher order oligomers. The C-terminal DNA-binding domains, linked to the scaffold by long flexible linkers, protrude into solution at regular intervals. This produces an array of DNA-binding domains. H-NS forms dimers with two-fold symmetry, sug-

gesting the higher order oligomers presents two arrays of DNA-binding domains, each located on opposite sides of the scaffold. These arrays of DNA-binding domains can then interact with stretches of DNA.

The model shown in Figure 5.1 is consistent with other studies of H-NS which highlight the cooperative nature of H-NS binding to DNA. With gel shift assays, the C-terminal domain of H-NS was estimated to bind DNA with a K_d of approximately 1 mM (Shindo *et al.*, 1995). Estimates of the affinity of wild-type H-NS for DNA (K_d) include 5 nM, determined by footprint titration experiments (Rimsky *et al.*, 2001) and 250 nM, determined using gel shift assays (Shindo *et al.*, 1995). Without a consensus of the DNA sequence and curvature to which H-NS optimally binds, comparison between the various estimates of the affinity of H-NS for DNA remains difficult. Nonetheless, it is apparent that the C-terminal domain in isolation binds DNA with affinities several orders of magnitude weaker than wild-type H-NS. Electrophoretic mobility shift assays which showed that H-NS_{1-64} does not interact with DNA is consistent with the 'lateral condensation' model, whereby the N-terminal domain of H-NS is not directly involved with DNA binding (Badaut *et al.*, 2002). Other studies have highlighted the importance of higher order oligomerisation with regards to H-NS function (see Section 1.9.6).

The model displays parallels with other studies of the H-NS/DNA interaction. Atomic force microscopy experiments have clearly shown lateral condensation of plasmid DNA, resulting in 'rod-like' structures (Dame *et al.*, 2000). Studies of the *virF* promoter suggest H-NS exerts its repressive effect by the formation of hairpin-like structures, where H-NS bridges together two lateral stretches of DNA (see Section 1.9.6 and Figure 1.10) (Prosseda *et al.*, 2004). This repressive hairpin-like structure has also been observed at the *rrnB* promoter (Dame *et al.*, 2002). The preference displayed by H-NS in binding curved DNA in preference to non-curved (see Section 1.9.5) may be rationalised the model shown in Figure 5.1, whereby the apex of a curved region of the DNA allows the stretches of DNA flanking that apex to come together and for the hairpin-like structure.

A precise figure for the stoichiometry of DNA-binding of H-NS (with respect to numbers of base pairs of DNA per molecule of H-NS) is yet to be determined. Esti-

mates of the stoichiometry vary between 15 to 20 bp of DNA per H-NS dimer (Amit *et al.*, 2003) to 14 bp of DNA per molecule of the C-terminal domain of H-NS (Shindo *et al.*, 1999).

The participation of H-NS in the mediation of an adaptive cellular response to a range of environmental stimuli, such as changes in oxygen availability, temperature, osmolarity and pH, is well documented (see Section 1.4). Efficient thermoregulatory mechanisms of pathogenic bacteria ensure expression of virulence factors occur only under the appropriate conditions. Several studies suggest alterations in DNA conformation in response to changes in the environmental conditions play a crucial role in some H-NS-dependent promoters (see Section 1.9.6). For example, a temperature-dependant change in conformation of DNA was shown to play a crucial role in the thermoregulation of the H-NS-dependant *virF* promoter (Prosseda *et al.*, 2004). However, the possibility that structural changes in H-NS modulate the expression of H-NS dependent genes has not been ruled out.

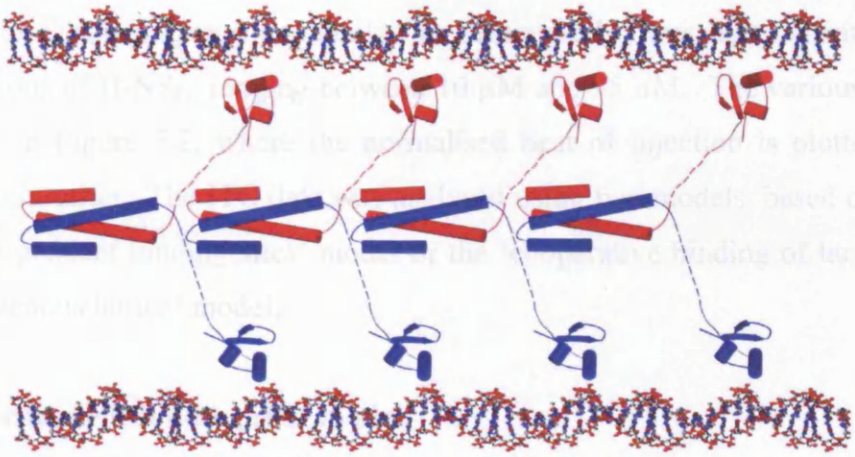


Figure 5.1
A schematic of the 'lateral condensation' model of the interaction between H-NS and DNA. H-NS oligomerises according to the head-to-tail model (see Section 1.9.4 and Chapter 4), presenting arrays of the C-terminal DNA-binding domains (Esposito *et al.*, 2002). These then bind two stretches of DNA, creating an inter-strand bridge.

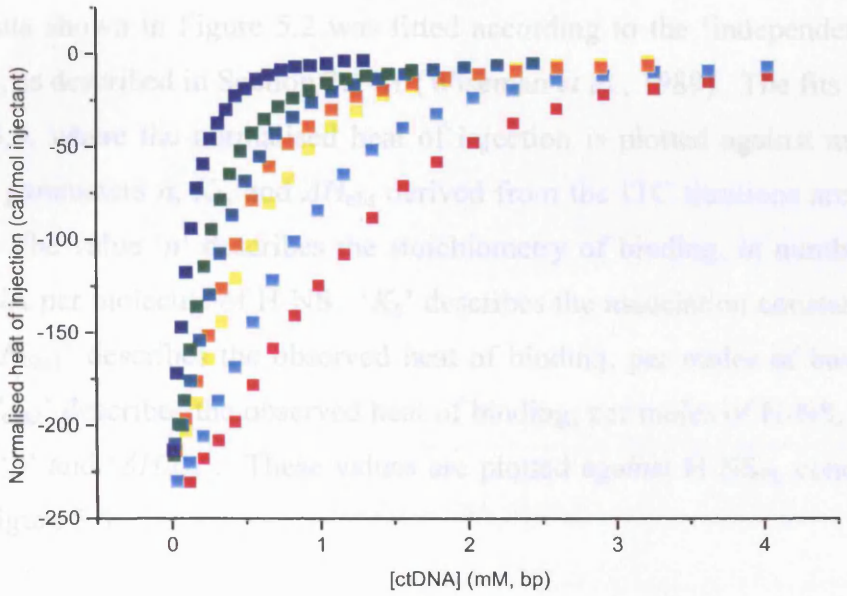


Figure 5.2
The various heats associated with ITC injections of DNA into H-NS_{FL}, normalised with respect to the DNA concentration, and plotted against DNA concentration. The following experiments are displayed: 8 mM DNA (bp) injected into 10 μM H-NS_{FL} (■), 10 mM DNA (bp) injected into 20 μM H-NS_{FL} (■), 15 mM DNA (bp) injected into 25 μM H-NS_{FL} (■), 20 mM DNA (bp) injected into 30 μM H-NS_{FL} (■), 20 mM DNA (bp) injected into 40 μM H-NS_{FL} (■), 25 mM DNA (bp) injected into 60 μM H-NS_{FL} (■), and 25 mM DNA (bp) injected into 95 μM H-NS_{FL} (■).

5.2 Examination of the interaction of H-NS with DNA using ITC

Calf thymus DNA, prepared as detailed in Section 2.14, was injected into varying concentrations of H-NS_{FL} ranging between 10 μM and 95 μM. The various titrations are shown in Figure 5.2, where the normalised heat of injection is plotted against DNA concentration. The ITC data was analysed using two models, based on binding to the 'independent binding sites' model or the 'cooperative binding of large ligands to a homogenous lattice' model.

5.2.1 'Independent binding site' model

The interaction of H-NS and DNA can be described using the simple model, termed the 'independent binding site' model, shown schematically in Figure 5.3. In this model the smallest binding unit of H-NS, the dimer, interacts with a stretch of DNA, n base pairs long, with association constant K_b. Hence a DNA molecule N base pairs long presents N/n independent binding sites. In this model, the cooperative effects of H-NS binding and the 'lattice effect' (described in Section 5.2.2) are disregarded.

The ITC data shown in Figure 5.2 was fitted according to the 'independent binding site' model, as described in Section 2.14.1 (Wiseman *et al.*, 1989). The fits are shown in Figure 5.4, where the normalised heat of injection is plotted against molar ratio. The fitting parameters n, K_b, and ΔH_{obs} derived from the ITC titrations are shown in Table 5.1. The value 'n' describes the stoichiometry of binding, in number of base pairs of DNA per molecule of H-NS. 'K_b' describes the association constant, in units of M^{-1}. 'ΔH_{obs1}' describes the observed heat of binding, per moles of base pairs of DNA. 'ΔH_{obs2}' describes the observed heat of binding, per moles of H-NS, and is the product of 'n' and 'ΔH_{obs1}'. These values are plotted against H-NS_{FL} concentration, shown in Figure 5.5.

It is interesting to note that whilst the stoichiometry appears to be independent of H-NS_{FL} concentration, K_b appears to have an inverse relationship with H-NS_{FL} concentration, suggesting H-NS binds DNA in an anti-cooperative fashion. However K_b appears to only exhibit a 4-fold difference over the range of H-NS_{FL} concentrations examined, and hence this relationship may not be significant. The heat of binding ΔH_{obs}, expressed in calories per mole of H-NS_{FL}, appears to be independent of H-NS_{FL} concentration.

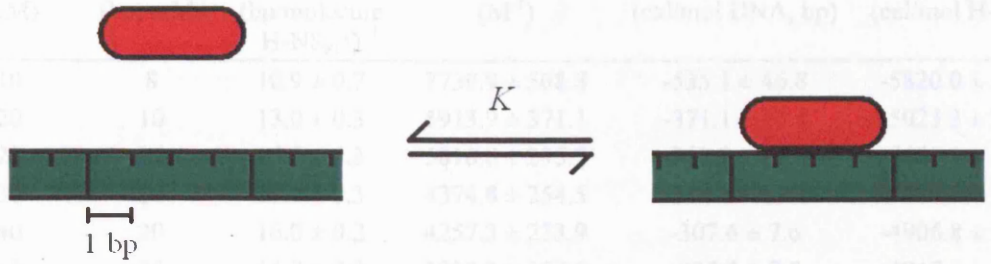


Figure 5.3
A schematic of the 'independent binding site model' describing the interaction between H-NS and DNA, with the H-NS dimer and DNA represented as red ovoids and green rectangles, respectively. For illustrative purposes only, the H-NS-binding site is shown as 4 bp long.

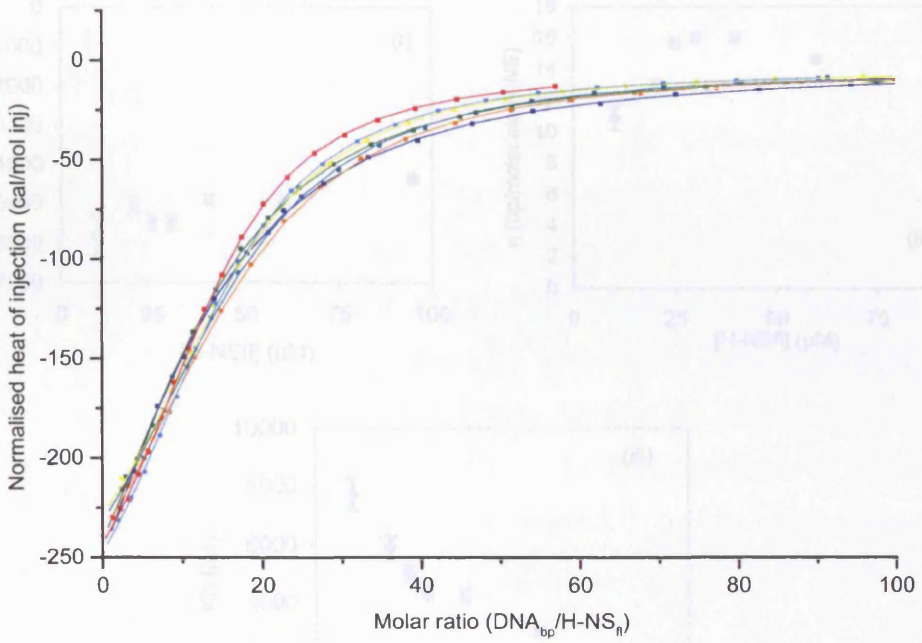


Figure 5.4
The various heats associated with ITC titrations of DNA into H-NS$_{FL}$, normalised with respect to DNA concentration, plotted against molar ratio (DNA, bp: H-NS$_{FL}$). The fits according to the 'independent binding sites' models are also shown. The following experiments are displayed: 8 mM DNA (bp) injected into 10 μM H-NS$_{FL}$ (■), 10 mM DNA (bp) injected into 20 μM H-NS$_{FL}$ (■), 15 mM DNA (bp) injected into 25 μM H-NS$_{FL}$ (■), 20 mM DNA (bp) injected into 30 μM H-NS$_{FL}$ (■), 20 mM DNA (bp) injected into 40 μM H-NS$_{FL}$ (■), 25 mM DNA (bp) injected into 60 μM H-NS$_{FL}$ (■), and 25 mM DNA (bp) injected into 95 μM H-NS$_{FL}$ (■).

[H-NS_{FL}] (μM)	[DNA] (bp, mM)	n (bp/molecule H-NS_{FL})	K_b (M^{-1})	ΔH_{obs1} (cal/mol DNA, bp)	ΔH_{obs2} (cal/mol H-NS_{FL})
10	8	10.9 ± 0.7	7739.9 ± 568.8	-535.1 ± 46.8	-5820.0 ± 509.1
20	10	13.0 ± 0.3	5913.9 ± 371.1	-371.1 ± 16.4	-5023.3 ± 212.5
25	15	15.7 ± 0.3	5010.8 ± 295.7	-349.2 ± 11.8	-5480.3 ± 185.7
30	20	16.1 ± 0.3	4374.8 ± 254.5	-342.0 ± 11.1	-5489.7 ± 178.1
40	20	16.0 ± 0.2	4257.3 ± 233.9	-307.6 ± 7.6	-4906.8 ± 121.6
60	25	14.7 ± 0.2	3000.2 ± 153.6	-335.3 ± 7.8	-4917.6 ± 113.8
95	25	13.9 ± 0.1	2369.9 ± 123.2	-313.5 ± 6.4	-4358.4 ± 88.9

Table 5.1
The fitting parameters n, K_b and ΔH_{obs} and the associated errors, derived from analysis of the various ITC titrations, as determined using the 'independent binding sites' model (Wiseman *et al.*, 1989).

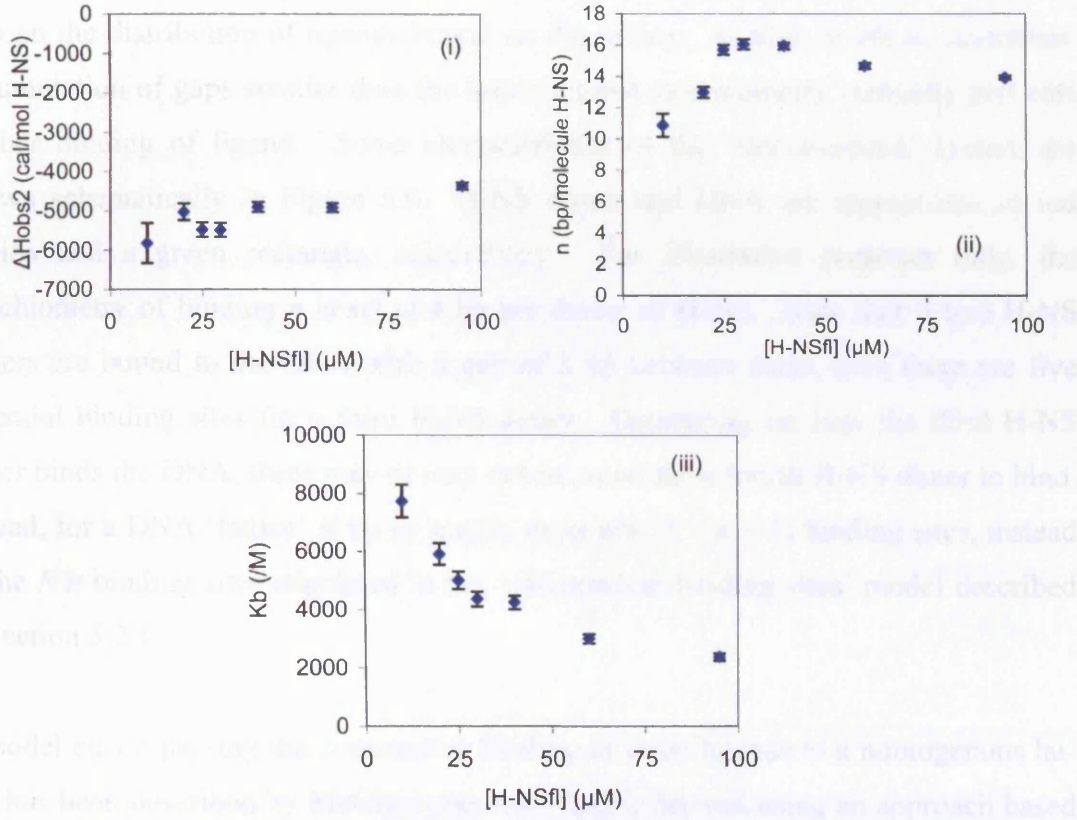


Figure 5.5
The values of ΔH_{obs} (cal/mol H-NS_{FL}), n, and K_b plotted against H-NS_{FL} concentration, determined using the 'identical binding sites' model, shown in (i), (ii) and (iii) respectively.

5.2.2 The model of cooperative binding of H-NS to a DNA lattice

DNA may be described as a one dimensional 'lattice' or an array of regularly repeating units. When considering a ligand that binds to the 'lattice' with little or no specificity whilst encompassing more than one unit of that 'lattice' when bound (in other words, the interaction between DNA and a protein that shows little or no sequence specificity whilst binding with a stoichiometry greater than 1 base pair per molecule of protein), a specific type of theoretical analysis is required (McGhee and von Hippel, 1974). These so-called 'lattice-ligand' systems include examples such as the binding of polymerases, histones, polyamines, antibiotics and dyes to DNA.

The main feature of lattice-ligand systems is that potential binding sites on the lattice overlap, due to the ligand protein encompassing more than one lattice residue when bound. Hence at any given level of saturation, the number of free ligand binding sites on the lattice depends not only on the amount of protein ligand already bound, but also on the distribution of ligands bound on the lattice. At high levels of saturation, accumulation of gaps smaller than the lattice-ligand stoichiometry sterically prevents further binding of ligand. Some characteristics of the 'lattice-ligand' system are shown schematically in Figure 5.6. H-NS dimer and DNA are represented as red ovoids and a green rectangle, respectively. For illustrative purposes only, the stoichiometry of binding n is set at 4 bp per dimer of H-NS. Note that if two H-NS dimers are bound to the DNA with a gap of 8 bp between them, then there are five potential binding sites for a third H-NS dimer. Depending on how the third H-NS dimer binds the DNA, there may or may not be room for a fourth H-NS dimer to bind. Indeed, for a DNA 'lattice' N bp in length, there are $(N - n + 1)$ binding sites, instead of the N/n binding sites stipulated in the 'independent binding sites' model described in Section 5.2.1.

A model encompassing the cooperative binding of large ligands to a homogenous lattice has been described by McGhee and von Hippel, derived using an approach based on conditional probabilities (McGhee and von Hippel, 1974). The plot of v/L against v is fitted to the following equation:

$$\frac{v}{L} = K(1-nv)\left(\frac{(2\omega+1)(1-nv)+v-R}{2(\omega-1)(1-nv)}\right)^{n-1}\left(\frac{1-(n+1)v+R}{2(1-nv)}\right)^{2} \qquad \text{Equation 2.11}$$

where v is the binding density (in moles of bound ligand per mole of lattice residue), L is the concentration of ligand (moles per litre), where n is the stoichiometry of binding and K_b is the intrinsic binding constant. The parameter ω describes cooperative binding: ω is defined as the ratio of the probability of ligands binding in a contiguous manner compared to the probability of ligands binding in an isolated manner. Values of $\omega < 1$, $\omega = 1$ and $\omega > 1$ describe anti-cooperative, non-cooperative and cooperative binding events, respectively. The binding of H-NS dimer to either an isolated, singly contiguous or doubly contiguous site is shown schematically in Figure 5.7.

The ITC data shown in Figure 5.2 was first rearranged to give the variables v and L which represent the binding density (moles of H-NS bound per mole DNA, bp) and DNA concentration (M, bp), respectively. The variables v and L are calculated from the raw ITC data using the method described in Section 2.14.2. The plot of v/L against v is shown in Figure 5.8, and was fitted according to Equation 2.11. The parameters n, ω and K_b, calculated from the fitting procedure, represent the stoichiometry, cooperativity factor and inherent binding constant respectively, and are detailed in Table 5.2. The parameters n, ω and K_b were plotted with respect to H-NS_{FL} concentration, as shown in Figure 5.9. All three parameters show inverse concentration dependence with respect to H-NS. These results are discussed in Section 5.4.

[H-NS_{FL}] (μM)	n	K_b (M^{-1})	ω
95	14.7 ± 0.4	569.0 ± 26.1	15.8 ± 2.1
60	18.5 ± 0.3	699.1 ± 5.7	35.5 ± 2.5
40	21.0 ± 0.2	888.1 ± 8.9	49.1 ± 2.3
30	23.3 ± 0.3	891.3 ± 9.8	61.5 ± 2.7
25	23.5 ± 0.3	1002.2 ± 11.2	65.5 ± 2.6
20	18.5 ± 0.3	1493.2 ± 4.6	35.3 ±1.4
10	23.9 ± 0.3	1730 ± 12.8	75.2 ± 2.0

Table 5.2
The parameters n, K_b, ω and their associated errors, as determined using the 'cooperative binding of large ligands to a homogenous lattice' model (McGhee and von Hippel, 1974)

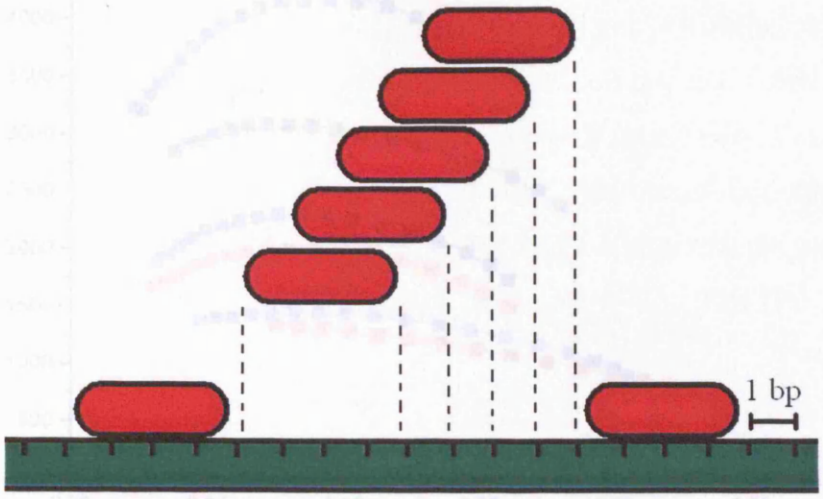


Figure 5.6
A schematic showing some characteristics of the 'lattice-ligand' interaction. H-NS dimers and DNA are represented as red ovoids and green rectangles, respectively. For illustrative purposes only, the stoichiometry of binding is shown as 4 bp per dimer of H-NS. For a gap of 8 bp, there are five potential H-NS binding sites. This figure was adapted from McGhee and von Hippel, 1974.

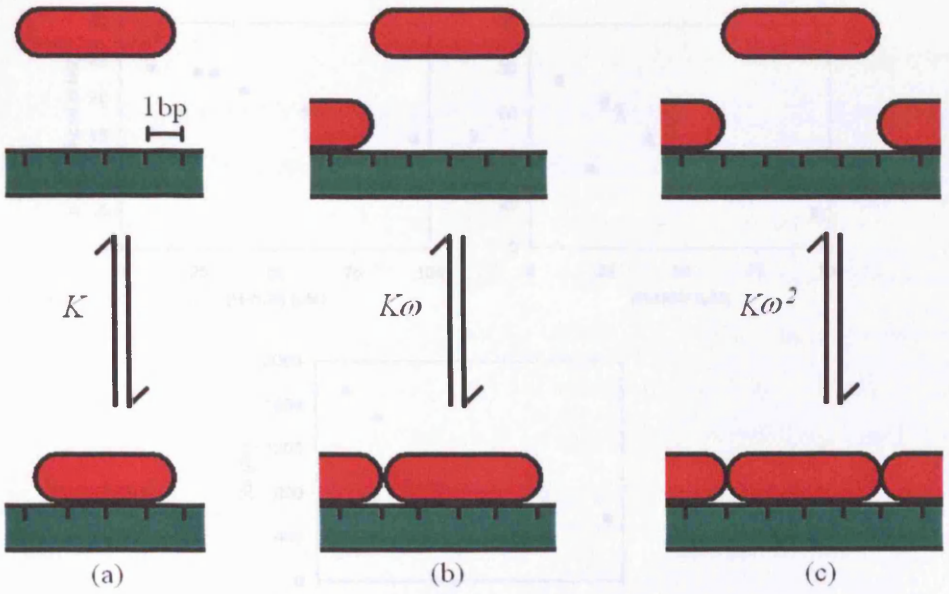


Figure 5.7
A schematic of the three distinct types of binding sites possible in 'lattice-ligand' systems, (a) 'isolated' (b) 'singly contiguous' and (c) 'doubly contiguous', and the respective associated binding constants. The ligand dimers of H-NS and DNA are represented as red ovoids and green rectangles, respectively. This figure was adapted from McGhee and von Hippel, 1974.

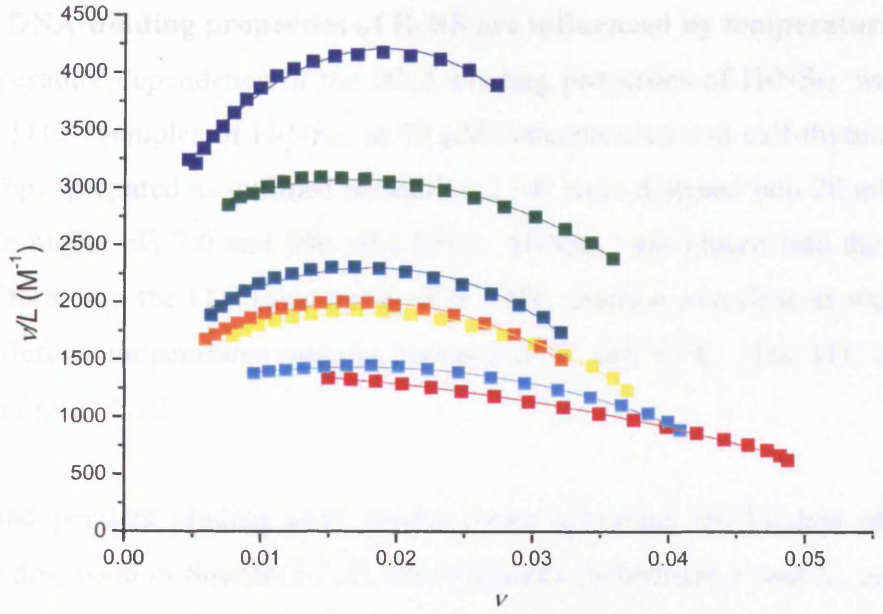


Figure 5.8
The plot of v/L against v of the various ITC titrations of H-NS_{FL} and DNA. The lines of best fit, are also shown, and the corresponding parameters shown in Table 5.2. The following experiments are displayed: 8 mM DNA (bp) injected into 10 μM H-NS_{FL} (■), 10 mM DNA (bp) injected into 20 μM H-NS_{FL} (■), 15 mM DNA (bp) injected into 25 μM H-NS_{FL} (■), 20 mM DNA (bp) injected into 30 μM H-NS_{FL} (■), 20 mM DNA (bp) injected into 40 μM H-NS_{FL} (■), 25 mM DNA (bp) injected into 60 μM H-NS_{FL} (■), and 25 mM DNA (bp) injected into 95 μM H-NS_{FL} (■).

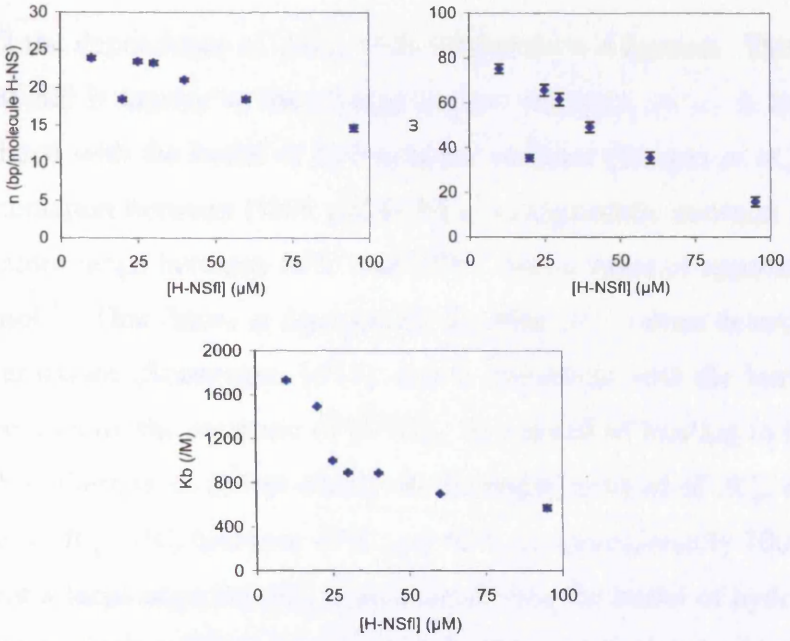


Figure 5.9
The values of n, ω and K_b describing the interaction of H-NS_{FL} and DNA, plotted with respect to H-NS_{FL} concentration, as determined using the 'lattice-ligand' model (McGhee and von Hippel, 1974). These values are derived from the fitting of the data shown in Figure 5.8 to Equation 2.11.

5.3 The DNA-binding properties of H-NS are influenced by temperature

The temperature dependence of the DNA-binding properties of H-NS_{FL} was investigated by ITC. Samples of H-NS_{FL} at 40 μM concentration and calf-thymus DNA at 20 mM (bp) (prepared as outlined in Section 2.14) were dialysed into 20 mM sodium phosphate buffer pH 7.0 and 300 mM NaCl. H-NS_{FL} was placed into the ITC cell, and the DNA into the ITC injection syringe. ITC titration experiments were carried out at different temperatures ranging between 10°C and 45°C. The ITC results are plotted in Figure 5.10

As the 'independent binding sites' model poorly describes the binding of H-NS to DNA (as discussed in Section 5.2.2), the calculated parameters n and K_b are of dubious relevance. Nonetheless, the fitting of the ITC data suggests the stoichiometry of binding is independent of temperature, with an average value of 15 base pairs of DNA per H-NS molecule. Whilst the association constant decreases with respect to temperature, only a four-fold difference in K_b is observed over the temperature range investigated (shown in Figure 5.11), and may not be significant. However, ITC measures the enthalpy directly and can be considered model-independent.

In Figure 5.12 the dependence of ΔH_{obs} with temperature is plotted. The temperature dependence of ΔH is known as the change in heat capacity, ΔC_p. A large negative ΔC_p is associated with the burial of hydrophobic surfaces (Cooper *et al.*, 2001). The ΔC_p of the interaction between DNA and H-NS_{FL} is apparently constant and negative in the temperature range between 10°C and 35°C, with a value of approximately -250 calories K^{-1} mol^{-1}. This figure is comparable to other ΔC_p values determined for bimolecular interactions (Sturtevant, 1977), and is consistent with the burial of hydrophobic surface area on the structure of H-NS_{FL} as a result of binding to DNA. However above 35°C there is an abrupt change in the negative trend of ΔC_p, resulting in a positive value of ΔC_p. ΔC_p between 40°C and 45°C is approximately 1000 calories K^{-1} mol^{-1}. Where a large negative ΔC_p is associated with the burial of hydrophobic surfaces, the converse is true, where a large positive ΔC_p is associated with the release of hydrophobic surfaces to solvent (Cooper *et al.*, 2001), and is generally observed upon the unfolding of a protein or a change in protein conformation resulting in the exposure of hydrophobic surfaces.

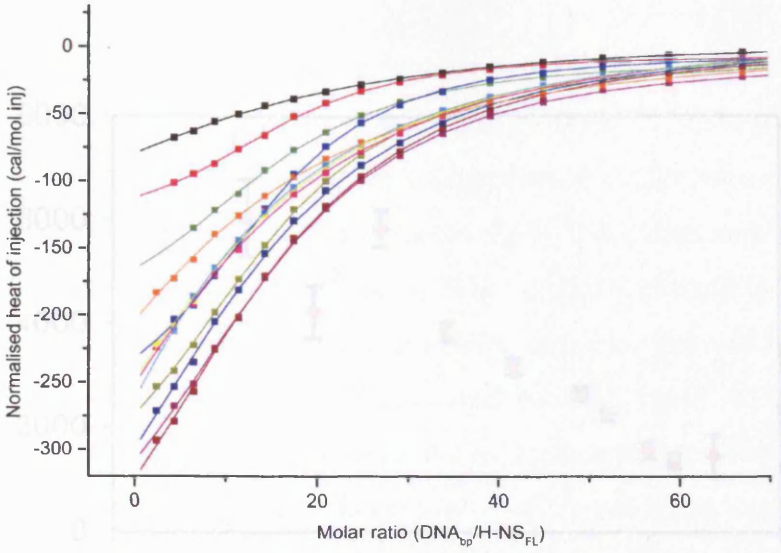


Figure 5.10
The various heats associated with ITC titrations of H-NS$_{FL}$ into DNA at different temperatures, normalised with respect to DNA, and plotted against molar ratio (DNA, bp: H-NS$_{FL}$), with fits according to the 'independent binding sites' model. H-NS$_{FL}$ was present in the ITC cell at 40 μM, and sonicated calf-thymus DNA was present in the injection needle at 20 mM (bp). ITC experiments were conducted at the following temperatures: 10°C (■), 15°C (■), 20°C (■), 25°C (■), 30°C (■), 35°C (■), 37°C (■), 40°C run A (■), 40°C run B (■), 40°C run B(■), 42°C (■), and 45°C (■). The binding parameters are summarised in Table 5.3

Temperature (°C)	n (bp/molecule H-NS)	K_b (M^{-1})	ΔH_{obs1} (cal/mol DNA, bp)	ΔH_{obs2} (cal/mol H-NS$_{FL}$)
10	16.6 ± 0.5	6034.4 ± 746.7	-132.7 ± 6.6	-2207.1 ± 125.0
15	15.9 ± 0.6	4231.0 ± 501.4	-215.0 ± 14.8	-3430.3 ± 272.7
20	15.5 ± 0.2	5843.5 ± 350.9	-291.1 ± 7.6	-4510.0 ± 135.2
25	15.7 ± 0.2	3847.4 ± 187.7	-379.2 ± 9.0	-5943.1 ± 162.4
30	16.2 ± 0.3	3188.4 ± 169.7	-447.4 ± 14.1	-7241.1 ± 264.8
35	15.5 ± 0.3	2668.9 ± 120.7	-523.6 ± 14.6	-8138.9 ± 262.5
37	14.4 ± 0.3	2256.6 ± 118.0	-542.5 ± 20.8	-7831.1 ± 351.4
40*	11.7 ± 1.9	1571.2 ± 166.4	-642.0 ± 126.0	-7333.5 ± 345.5
42	12.5 ± 1.0	1357.6 ± 137.3	-507.0 ± 60.0	-6344.8 ± 911.7
45	14.9 ± 2.8	1457.4 ± 400.1	-178.5 ± 50.2	-2666.1 ± 899.3

Table 5.3
The fitting parameters n, K_b, ΔH_{obs1} and ΔH_{obs2}, derived from analysis of the various ITC titrations at different temperatures, as determined using the 'independent binding sites' (Wiseman *et al.*, 1989). The fittings for all of the data points (except where a temperature of 25 °C is used) have been derived from single experiments. The errors associated with all the parameters are the fitting errors.
* The fittings for this data point are derived from the average from three identical readings. The associated errors have been calculated assuming the parameters follow a normal distribution.

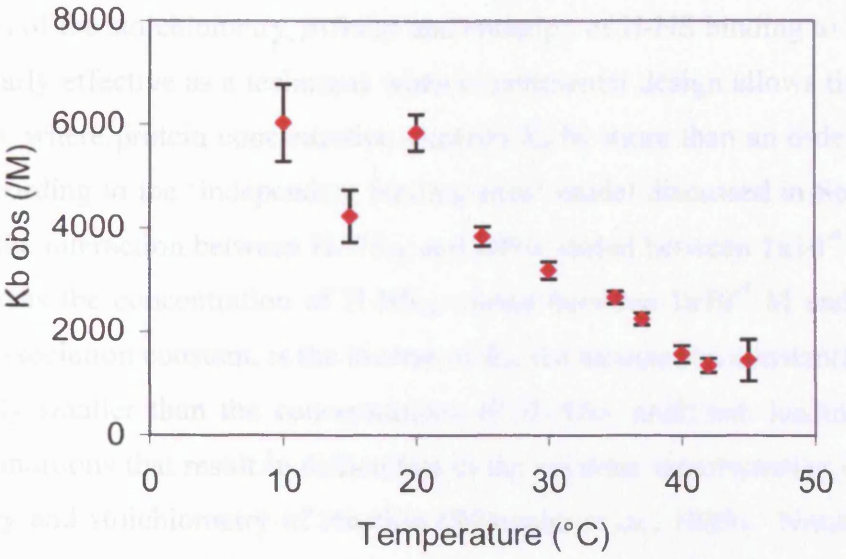


Figure 5.11
The dependence of K_b with temperature of the association of H-NS_{FL} with DNA. The values were derived from the fitting of ITC data with the 'independent binding sites' model (Wiseman *et al.*, 1989). The data is taken from Table 5.3

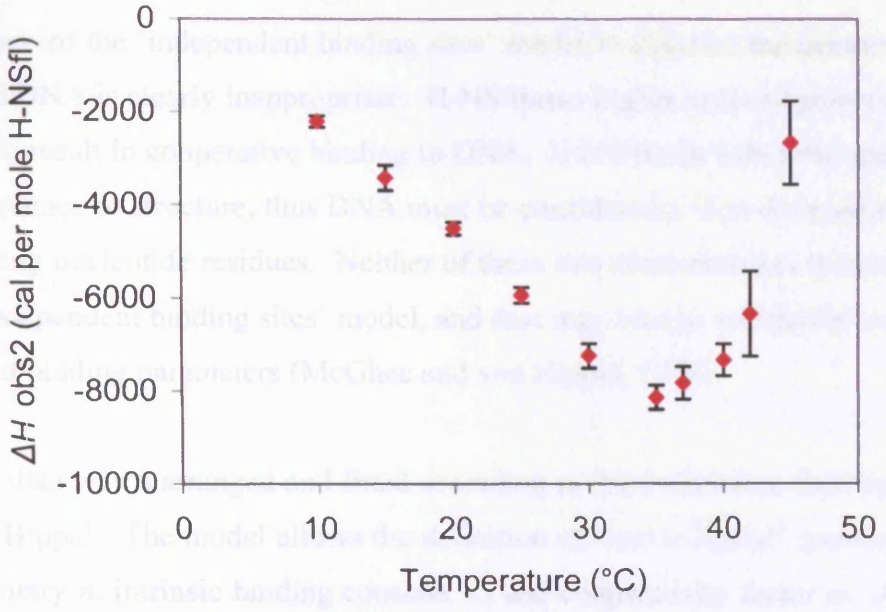


Figure 5.12
The temperature dependence of ΔH_{obs} of the association of H-NS_{FL} and DNA. The values were derived from the fitting of ITC data with the 'independent binding sites' model (Wiseman *et al.*, 1989). The data is taken from Table 5.3. The gradient of the plot yields ΔC_p, the change in heat capacity.

5.4 Discussion

The interaction of H-NS_{FL} and DNA was investigated with ITC in order to obtain precise values of the stoichiometry, affinity and enthalpy of H-NS binding to DNA. ITC is particularly effective as a technique when experimental design allows tight-binding conditions, where protein concentration exceeds K_d by more than an order of magnitude. According to the 'independent binding sites' model discussed in Section 5.2.1, the K_d of the interaction between H-NS_{FL} and DNA varied between $1x10^{-4}$ M to $4x10^{-4}$ M, whereas the concentration of H-NS_{FL} varied between $1x10^{-5}$ M and $1x10^{-4}$ M. (K_d, the dissociation constant, is the inverse of K_b, the association constant). The K_d is consistently smaller than the concentrations of H-NS_{FL} analysed, leading to weak-binding conditions that result in difficulties in the accurate determination of the binding affinity and stoichiometry of reaction (Wiseman *et al.*, 1989). Nonetheless, the ITC data revealed several features of the interaction between H-NS_{FL} and DNA. The stoichiometry of binding averaged at 14.3 moles of base pairs of DNA per mole of H-NS_{FL}. Interestingly the apparent binding affinity varied inversely with respect to H-NS_{FL} concentration, suggesting H-NS_{FL} binds DNA in an anti-cooperative manner. However this is clearly at odds with the current knowledge of H-NS function.

Application of the 'independent binding sites' model to describe the interaction of H-NS_{FL} and DNA is clearly inappropriate. H-NS forms higher order oligomers, which is thought to result in cooperative binding to DNA. H-NS binds with little specificity to DNA sequence or structure, thus DNA must be considered a 'one-dimensional lattice' of repeating nucleotide residues. Neither of these two characteristics is accounted for in the 'independent binding sites' model, and thus may lead to substantial errors in the calculated binding parameters (McGhee and von Hippel, 1974).

The ITC data was rearranged and fitted according to the model described by McGhee and von Hippel. The model allows the definition of 'lattice-ligand' systems with the stoichiometry n, intrinsic binding constant K_b and cooperativity factor ω. A description of these parameters is detailed in Section 5.2.2. The three parameters displayed inverse relationships with H-NS_{FL} concentration, suggesting the model may not describe adequately the interaction between H-NS_{FL} and DNA. If the model satisfactorily describes the interaction between H-NS_{FL} and DNA, the parameters should be invariant with respect to H-NS_{FL} concentration under any specific set on conditions.

A variety of possible reasons may invalidate the model described by McGhee and von Hippel. According to the 'lateral condensation' model for the interaction between H-NS and DNA, the oligomers of H-NS themselves may be considered 'lattices', a characteristic that is not accounted for in the model described by McGhee and von Hippel. Thus the H-NS/DNA interaction may be described as 'lattice-lattice' interactions as opposed to 'lattice-ligand'. There may be many characteristics of the interaction between H-NS and DNA that invalidate the use of the two models describe in this chapter. For instance, mounting evidence suggests that H-NS can bind DNA in two distinguishable ways. Fluorospectroscopic studies have identified different modes of interaction of H-NS with DNA, depending on whether the interaction is specific or not (Tippner and Wagner, 1995), suggesting that there are at least two modes of interaction (and hence two binding affinities). Atomic force microscopy experiments have shown that H-NS condenses lateral stretches of DNA, leading to compaction (Dame *et al.*, 2001). It is not known in the experiments described in this chapter whether H-NS will condense two DNA molecules together, or whether H-NS will cause one DNA molecule to fold back on itself to form a hairpin-like structure. H-NS has been shown to display differential effects on DNA supercoiling depending on the amount of H-NS relative to DNA (Tupper *et al.*, 1994). At saturating concentrations, H-NS has been shown to increase the bending rigidity and persistence length of λ-phage DNA, in contrast to other studies involving sub-saturating H-NS concentrations which suggest DNA is compacted upon binding of H-NS (Amit *et al.*, 2003; Dame and Wuite, 2003). Clearly, in an ITC experiment, where conditions shift from excess of H-NS to excess of DNA, the changing relative levels of H-NS and DNA will affect the interaction.

The interaction of DNA and $H\text{-}NS_{FL}$ at differing temperatures was examined by ITC. The binding stoichiometry was found to be 15 base pairs of DNA per $H\text{-}NS_{FL}$ molecule, and independent of temperature. The association constant K_b was found to vary inversely with respect to temperature, approximately four-fold over the range tested. However, the suitability of the use of the model used to fit the ITC data, the 'independent binding sites' model (Wiseman *et al.*, 1989), is questionable. On the other hand, determination of ΔH can be considered model-independent. The ITC data suggested a strong dependence of ΔH with respect to temperature, the slope of which gives the change in heat capacity, ΔC_p. Between the temperatures 10°C to 35°C there

was a constant negative slope. The corresponding ΔC_p value was consistent with the burial of hydrophobic surface area on H-NS_{FL} as a result of binding DNA. However between 37°C and 45°C the slope of the plot changes abruptly, resulting in a positive ΔC_p. This suggests that H-NS_{FL} either is unfolding above 37°C, or is undergoing a conformational exchange, which can account for the net exposure of hydrophobic surfaces above 37°C. A variety of biophysical studies suggests that H-NS_{FL} has a relatively high melting temperature. For example, NMR melting experiments suggest an apparent melting point of 58°C for H-NS_{FL} (Smyth *et al.*, 2000). CD measurements of thermal denaturation of H-NS_{FL} at 288 nm suggested a melting temperature of 58°C, whilst several DSC scans of H-NS_{FL} revealed an average melting temperature of 52°C (Smyth, 1999). However, one CD study of the thermal denaturation at 220 nm of H-NS_{FL} suggested the melting temperature of 40.81°C (Schroder *et al.*, 2001). On the weight of this evidence, H-NS_{FL} is unlikely to exhibit unfolding between 10°C to 45°C; instead the ITC data implies that H-NS_{FL} undergoes a conformational change at around 35°C. The implications of the ITC data on the thermoregulatory functions of H-NS are discussed in further detail in Chapter 6.

Chapter 6

Conclusions

Roles for H-NS and its related proteins have been implicated in a diverse range of physiological functions within the bacterial cell, mainly at the level of repression of transcription of H-NS-dependent genes. H-NS modulates expression of these genes in response to a variety of stimuli (see Section 1.4), thus can be viewed as an integral part of the mechanism by which the cell adapts to changing environmental challenges. Consistent with that role, the importance of H-NS has been highlighted in the modulation of virulence factors in pathogenic organisms.

An understanding of the structure-function relationship of H-NS has started to emerge from recent experiments. The 'head-to-tail' model of the oligomerisation of H-NS (see Section 1.9.4) best incorporates the two key phenomena of oligomerisation and DNA-binding, and is consistent with many biophysical studies to date. According to this model, the N-terminal domain oligomerises in a 'head-to-tail' fashion, thereby resulting in a protein 'scaffold' that presents DNA-binding domains in a regular array. The experiments described in Chapter 4 are consistent with the 'head-to-tail' mechanism, as well as providing an idea of the time scale over which these interactions occur. The exchange of the monomers of the dimeric unit of H-NS has also been inves-

tigated, and the implications to the 'head-to-tail' model discussed (see Section 4.9). These experiments provide a structural basis of the dominant-negative effect observed with previously reported H-NS mutants (Williams *et al.*, 1996; Ueguchi *et al.*, 1996).

The 'head-to-tail' model is incorporated into the 'lateral condensation model', where the higher order oligomers of H-NS binds to two stretches of DNA, creating an inter-strand bridge and condensing them (see Section 5.1). Lateral condensation has been observed with atomic force microscopy of H-NS/DNA complexes, and is consistent with the repressive hairpin-like nucleoprotein structures observed at the *virF* and *rrnB* P1 promoters (Dame *et al.*, 2000; Dame *et al.*, 2002; Prosseda *et al.*, 2004). The preference in binding for curved over non-curved DNA exhibited by H-NS (see Section 1.9.5) infers that the modulation of expression of H-NS-dependent genes through the formation of these hairpin-like structures represents a general mechanism of H-NS function. However a further level of complexity in H-NS function is added by the requirement of specific regulatory proteins required at different promoters (see Section 1.7).

In line with the role of H-NS in environmental adaptation, the nucleoprotein structure of H-NS and DNA appears responsive to environmental stimuli. For instance, changes in temperature and osmolarity lead to differences in DNA-binding properties of H-NS (Amit *et al.*, 2003). Clear temperature-dependent changes in the oligomeric properties of $H\text{-}NS_{1\text{-}89}$ are observed by SEC (see Section 4.2). Analysis of the interaction of H-NS and DNA at differing temperatures using ITC suggest a conformational change in H-NS occurs at around 35 °C (see Section 5.3).

There are two model NMR structures of the N-terminal domain of H-NS (Esposito *et al.*, 2002; Bloch *et al.*, 2003), with one model determined at 25°C and the other at 35°C. On the basis of the ITC data it is tempting to speculate that these two model structures represent two different temperature-dependent conformations. Interestingly, the differences observed in the expression of many genes under thermoregulatory control by H-NS occur between 25°C and 37°C (Soutourina *et al.*, 2002; Umanski *et al.*, 2002; Poore and Mobley, 2003; Prosseda *et al.*, 2004). If different functional properties can be attributed to the two conformers of H-NS, for instance differential oligomerisation properties or differential DNA-binding properties (though these two

properties of H-NS are not mutually exclusive), then the mechanism by which H-NS exerts its thermoregulatory control may be rationalised through a temperature-dependent structural change at around 35°C. Indeed SEC traces of H-NS_{1-89} between 17.5 °C and 45 °C (see Section 4.2 and Figure 4.1) show significant differences in the oligomerisation properties in this temperature range, suggesting thermoregulation may be effected at the protein level.

There are matters of concern associated with this hypothesis. Firstly, the thermoregulatory control of gene expression from the *virF* promoter has been shown to be mediated through a highly specific temperature-dependent change in the conformation of the promoter DNA (Prosseda *et al.*, 2004), although a temperature-dependent conformational change in H-NS has not been ruled out. Other studies highlight the importance of changes in DNA conformation leading to changes in gene expression, as a result of a change in environmental conditions (Ueguchi and Mizuno, 1993; Jordi and Higgins, 2000). Indeed DNA topology has been shown to depend on environmental parameters such as osmolarity, anoxia, nutrient shifts, growth phase, pH and temperature (Drlica, 1992; Karem and Foster, 1993). It is likely that the effects of H-NS depend on the target DNA promoter, thus allowing more specific regulatory control over the gene expression of any given gene. Indeed, in terms of evolutionary changes in the response of the cell to environmental stimuli, any mutation in H-NS may result in a global change in regulatory control of H-NS-dependent genes. On the other hand, mutations in DNA sequence of individual H-NS-dependent promoters (and hence the sensitivity of DNA conformation of that promoter to environmental stimuli) would allow specific changes in the cell without disrupting the other functions of H-NS. Another issue associated with the proposal of a temperature-dependent conformational switch of H-NS effecting thermoregulatory control of gene expression is the difficulty of extending this model to instances where H-NS regulates gene expression in response to changes in environment other than temperature. The suggested temperature-dependent conformational switch is unlikely to occur in changes in a diversity of conditions such as anoxia, nutrient shift, pH and osmolarity over a physiologically relevant range. Nevertheless it is clear that changes in temperature lead to observable differences in the oligomeric properties and DNA-binding properties of H-NS, thereby providing an insight into the mechanism of thermoregulation of genes under H-NS control.

Precise modulation of expression of different H-NS-dependent promoters may be mediated by specific regulatory proteins including H-NS. Varying ratios of these proteins (e.g. H-NS, StpA and Sfh in Shigella flexneri 2a strain 2457T) may result in differential control of cellular processes (Dorman and Deighan, 2003). With 58 % sequence identity with H-NS, the paralogous protein H-NS presents an opportunity of study of the role of auxiliary proteins in the H-NS-mediated repression of gene expression. A specific role for StpA has yet to be defined (see Section 1.11), though it is clear that it assumes a minor role with respect to H-NS (see Section 1.12). Several interesting features of H-NS and StpA were identified by comparing and contrasting with a series of biophysical experiments.

Model structures of the C-terminal domain of StpA were determined by NMR (see Chapter 3). Of the two methods used, termed the manual and automatic methods, the latter was shown to result in a structure of higher quality and without potential bias. $StpA_{91-134}$ forms a small, loose, solvent exposed structure with two α-helices and two β-strands. Unfortunately, the structural determinant on $StpA_{91-134}$ involved DNA-binding could not be identified. The structure of the C-terminal domain of StpA was shown to be very similar to that of H-NS: the secondary structure elements and tertiary fold of the two domains aligned very well with each other. However the C-terminal domain of H-NS was suggested to be far more flexible than that of StpA. As to whether the C-terminal domain of H-NS truly represents a more inherently flexible structure than the equivalent of StpA cannot be determined, due to the very small number of restraints used to determine the former structure.

StpA was shown to have a similar domain organisation to H-NS (see Chapter 4). SEC experiments suggest the formation of higher order oligomers of StpA is mediated by the N-terminal domain in a 'head-to-tail' fashion analogous to H-NS. NMR experiments suggest the C-terminal domain is linked to the N-terminal domain by a long flexible linker, and exhibits different mobilities in solution compared to the N-terminal domain. CD experiments and *in silico* analysis suggest that the N-terminal domain of StpA forms an α-helix-rich domain with a coiled-coil region between residues 20 to 60. Interestingly, the CD melting studies indicated a difference in thermal stability between H-NS and StpA. Other clear differences were observed between H-

NS and StpA in their oligomerisation properties. Both the monomer exchange between homodimers and the formation of homooligomers of StpA apparently occur with higher affinity and on much faster time scale than H-NS.

The SEC experiments described in Chapter 4 provide many insights into the interaction between H-NS and StpA. The pull-down assays and the SEC experiments show that StpA can interact with H-NS through both the coiled-coil interface in the N-terminal domain and through the formation of higher order heterooligomers. Whilst the resolution of the SEC experiments is insufficient to differentiate between the two phenomena, it is reasonable to conclude that both types of interaction occur. Crucially, these experiments showed that StpA interacts with H-NS rapidly and with high affinity. The affinity of H-NS and StpA for each other and the speed of that interaction (whether through the coiled-coil interface or though the 'head' and 'tail' oligomerisation interfaces) are apparently higher than the equivalent homomeric interactions. This is consistent with studies that had shown H-NS to confer resistance to StpA to Lon protease through the formation of heterooligomers (Johansson and Uhlin, 1999; Johansson *et al.*, 2001).

On the basis of available evidence, a model of the *in vivo* role of StpA can be proposed. Clearly H-NS and StpA function are linked. The autoregulation of the *hns* promoter is linked to DNA-synthesis, so that there is a relatively constant ratio of H-NS to DNA (Free and Dorman, 1995), and the expression of H-NS is relatively insensitive to environmental challenges (see Section 1.6), suggesting H-NS mediates the maintenance of intracellular homeostasis. On the other hand, the expression of the *stpA* promoter is strongly repressed by H-NS, and is responsive to changes in environmental conditions (see Section 1.12). (In fact, H-NS and StpA display auto- and cross-regulation at the *hns* and *stpA* promoters.) As the experiments described in Chapter 4 suggest, any free StpA (i.e. as soon as it is expressed in the cell) is immediately incorporated into the highly abundant higher order oligomers of H-NS present in the cell, otherwise the free StpA is rapidly degraded by Lon protease (Johansson and Uhlin, 1999). StpA displays different DNA- and RNA-binding properties compared to H-NS (see Sections 1.14 and 1.15). Thus the DNA-binding properties of the higher order heterooligomeric complexes of H-NS and StpA in the bacterial cell (consisting predominantly of H-NS) can be fine-tuned by the ratio of H-NS to StpA. In turn, this

ratio can be manipulated by the combination of expression from the environmentally sensitive promoter of *stpA* and the rapid degradation of free StpA by Lon. There may be no unique function of StpA other than to modify the characteristics of the higher order oligomers of H-NS, suggesting a specific role of StpA is unlikely to be identified.

References

Afflerbach, H., Schroder, O., and Wagner, R. (1999) Conformational changes of the upstream DNA mediated by H-NS and FIS regulate *E. coli* RrnB P1 promoter activity. *J. Mol. Biol.* **286**, 339-353.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* **181**, 6361-70.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.

Amit, R., Oppenheim, A. B., and Stavans, J. (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys. J* **84**, 2467-2473.

Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) An alternative 3D NMR technique for correlating backbone ^{15}N with side-chain H resonances in larger proteins. *J Magn Reson* **95**, 636-641.

Atlung, T. and Ingmer, H. (1997) H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* **24**, 7-17.

Azam, T. A., Hiraga, S., and Ishihama, A. (2000) Two types of localization of the DNA-binding proteins within the *Escherichia coli* nucleoid. *Genes Cells* **5**, 613-26.

Azam, T. A. and Ishihama, A. (1999) Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**, 33105-13.

Badaut, C., Williams, R., Arluison, V., Bouffartigues, E., Robert, B., Buc, H., and Rimsky, S. (2002) The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *J Biol. Chem* **277**, 41657-41666.

Bang, I. S., Audia, J. P., Park, Y. K., and Foster, J. W. (2002) Autoinduction of the *ompR* response regulator by acid shock and control of the *Salmonella enterica* acid tolerance response. *Mol Microbiol.* **44**, 1235-1250.

Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**, 5269-5278.

Bax, A. (1989) Two-dimensional NMR and protein structure. *Annu. Rev. Biochem.* **58**, 223-256.

Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) Measurement of homo- and heteronuclear *J* couplings from quantitative *J* correlation. *Methods Enzymol.* **239**, 79-105.

Beloin, C., Deighan, P., Doyle, M., and Dorman, C. J. (2003) *Shigella flexneri* 2a strain 2457T expresses three members of the H-NS-like protein family: characterization of the Sfh protein. *Mol Genet. Genomics* **270**, 66-77.

Beloin, C. and Dorman, C. J. (2003) An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of *Shigella flexneri*. *Mol Microbiol.* **47**, 825-838.

Bertin, P., Benhabiles, N., Krin, E., Laurent-Winter, C., Tendeng, C., Turlin, E., Thomas, A., Danchin, A., and Brasseur, R. (1999) The structural and functional organization of H-NS-like proteins is evolutionarily conserved in gram-negative bacteria. *Mol Microbiol* **31**, 319-29.

Bertin, P., Hommais, F., Krin, E., Soutourina, O., Tendeng, C., Derzelle, S., and Danchin, A. (2001) H-NS and H-NS-like proteins in Gram-negative bacteria and their multiple role in the regulation of bacterial metabolism. *Biochimie* **83**, 235-41.

Bloch, V., Yang, Y., Margeat, E., Chavanieu, A., Auge, M. T., Robert, B., Arold, S., Rimsky, S., and Kochoyan, M. (2003) The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat. Struct. Biol.* **10**, 212-218.

Boucher, W. (2002) AZARA version 2.7.

Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Cryst.* **D54**, 905-921.

Buggy, J. J., Sganga, M. W., and Bauer, C. E. (1994) Characterization of a light-responding *trans*-activator responsible for differentially controlling reaction center and light-harvesting-I gene expression in *Rhodobacter capsulatus*. *J Bacteriol.* **176**, 6936-6943.

Bustamante, V. H., Santana, F. J., Calva, E., and Puente, J. L. (2001) Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. *Mol. Microbiol.* **39**, 664-678.

Cerdan, R., Bloch, V., Yang, Y., Bertin, P., Dumas, C., Rimsky, S., Kochoyan, M., and Arold, S. T. (2003) Crystal structure of the N-terminal dimerisation domain of VicH, the H-NS-like protein of *Vibrio cholerae*. *J Mol Biol.* **334**, 179-185.

Ceschini, S., Lupidi, G., Coletta, M., Pon, C. L., Fioretti, E., and Angeletti, M. (2000) Multimeric self-assembly equilibria involving the histone-like protein H-NS. A thermodynamic study. *J Biol Chem* **275**, 729-34.

Cooper, A., Johnson, C. M., Lakey, J. H., and Nollmann, M. (2001) Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophys. Chem.* **93**, 215-230.

Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol. NMR* **13**, 289-302.

Cusick, M. E. and Belfort, M. (1998) Domain structure and RNA annealing activity of the *Escherichia coli* regulatory protein StpA. *Mol Microbiol* **28**, 847-57.

Dame, R. T. and Goosen, N. (2002) HU: promoting or counteracting DNA compaction? *FEBS Lett.* **529**, 151-156.

Dame, R. T. and Wuite, G. J. (2003) On the role of H-NS in the organization of bacterial chromatin: from bulk to single molecules and back. *Biophys. J* **85**, 4146-4148.

Dame, R. T., Wyman, C., and Goosen, N. (2000) H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* **28**, 3504-3510.

Dame, R. T., Wyman, C., and Goosen, N. (2001) Structural basis for preferential binding of H-NS to curved DNA. *Biochimie* **83**, 231-234.

Dame, R. T., Wyman, C., Wurm, R., Wagner, R., and Goosen, N. (2002) Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the *rrnB P1*. *J Biol Chem* **277**, 2146-50.

Deighan, P., Beloin, C., and Dorman, C. J. (2003) Three-way interactions among the Sfh, StpA and H-NS nucleoid-structuring proteins of *Shigella flexneri* 2a strain 2457T. *Mol Microbiol.* **48**, 1401-1416.

Deighan, P., Free, A., and Dorman, C. J. (2000) A role for the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of *micF* RNA stability. *Mol Microbiol* **38**, 126-39.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol. NMR* **6**, 277-293.

Delano, W. L. (2004) The PyMOL Molecular Graphics System.

Delihas, N. and Forst, S. (2001) MicF: an antisense RNA gene involved in response of *Escherichia coli* to global stress factors. *J Mol Biol* **313**, 1-12.

Dersch, P., Kneip, S., and Bremer, E. (1994) The nucleoid-associated DNA-binding protein H-NS is required for the efficient adaptation of *Escherichia coli* K-12 to a cold environment. *Mol Gen Genet* **245**, 255-9.

Dersch, P., Schmidt, K., and Bremer, E. (1993) Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol Microbiol.* **8**, 875-889.

Dole, S., Nagarajavel, V., and Schnetz, K. (2004) The histone-like nucleoid structuring protein H-NS represses the *Escherichia coli bgl* operon downstream of the promoter. *Mol Microbiol.* **52**, 589-600.

Doreleijers, J. F., Raves, M. L., Rullmann, J. A. C., and Kaptein, R. (1999) Completeness of NOEs in proteins: a statistical analysis of NMR data. *J Biomol. NMR* **14**, 123-132.

Dorman, C. J. (2004) H-NS: a universal regulator for a dynamic genome. *Nat. Rev. Microbiol.* **2**, 391-400.

Dorman, C. J. and Deighan, P. (2003) Regulation of gene expression by histone-like proteins in bacteria. *Curr. Opin. Genet. Dev.* **13**, 179-184.

Dorman, C. J., Hinton, J. C., and Free, A. (1999) Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol* **7**, 124-8.

Dorman, C. J., Porter, M. E. (1998) The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. *Mol. Microbiol.* **29**, 677-684.

Drlica, K. (1992) Control of bacterial DNA supercoiling. *Mol. Microbiol.* **6**, 425-433.

Esposito, D., Petrovic, A., Harris, R., Ono, S., Eccleston, J. F., Mbabaali, A., Haq, I., Higgins, C. F., Hinton, J. C., Driscoll, P. C., and Ladbury, J. E. (2002) H-NS oligomerisation domain structure reveals the mechanism for high order self-association of the intact protein. *J Mol. Biol.* **324**, 841-850.

Falconi, M., Brandi, A., La Teana, A., Gualerzi, C. O., and Pon, C. L. (1996) Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Mol Microbiol.* **19**, 965-975.

Falconi, M., Colonna, B., Prosseda, G., Micheli, G., and Gualerzi, C. O. (1998a) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A tempera-

ture-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *Embo J* **17**, 7033-7043.

Falconi, M., Higgins, N. P., Spurio, R., Pon, C. L., and Gualerzi, C. O. (1993) Expression of the gene encoding the major bacterial nucleotide protein H-NS is subject to transcriptional auto-repression. *Mol Microbiol.* **10**, 273-282.

Falconi, M., Prosseda, G., Giangrossi, M., Beghetto, E., and Colonna, B. (2001) Involvement of FIS in the H-NS-mediated regulation of *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. *Mol. Microbiol.* **42**, 439-452.

Free, A. and Dorman, C. J. (1995) Coupling of *Escherichia coli hns* mRNA levels to DNA synthesis by autoregulation: implications for growth phase control. *Mol Microbiol.* **18**, 101-113.

Free, A. and Dorman, C. J. (1997) The *Escherichia coli stpA* gene is transiently expressed during growth in rich medium and is induced in minimal medium and by stress conditions. *J Bacteriol.* **179**, 909-918.

Free, A., Porter, M. E., Deighan, P., and Dorman, C. J. (2001) Requirement for the molecular adapter function of StpA at the *Escherichia coli bgl* promoter depends upon the level of truncated H-NS protein. *Mol Microbiol.* **42**, 903-17.

Free, A., Williams, R. M., and Dorman, C. J. (1998) The StpA protein functions as a molecular adapter to mediate repression of the *bgl* operon by truncated H-NS in *Escherichia coli*. *J Bacteriol* **180**, 994-7.

Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.

Goransson, M., Sonden, B., Nilsson, P., Dagberg, B., Forsman, K., Emanuelsson, K., and Uhlin, B. E. (1990) Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* **344**, 682-685.

Goyard, S. and Bertin, P. (1997) Characterization of BpH3, an H-NS-like protein in *Bordetella pertussis*. *Mol Microbiol* **24**, 815-23.

Grzesiek, S. and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. *J Biomol. NMR* **3**, 185-204.

Haack, K. R., Robinson, C. L., Miller, K. J., Fowlkes, J. W., and Mellies, J. L. (2003) Interaction of Ler at the LEE5 (*tir*) operon of enteropathogenic *Escherichia coli*. *Infect. Immun.* **71**, 384-392.

Hagemann, M., Schoor, A., Jeanjean, R., Zuther, E., and Joset, F. (1997) The *stpA* gene from *Synechocystis* sp. strain PCC 6803 encodes the glucosylglycerol-phosphate phosphatase involved in cyanobacterial osmotic response to salt shock. *J. Bacteriol.* **179**, 1727-1733.

Hall, T. (2004) BioEdit Sequence Alignment Editor v6.0.7.

Hinton, J. C. (1997) The *Escherichia coli* genome sequence: the end of an era or the start of the FUN? *Mol Microbiol* **26**, 417-22.

Hirschbein, L. and Guillen, N. (1982) Characterization, assay, and use of isolated bacterial nucleoids. *Methods Biochem. Anal.* **28**, 297-328.

Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J. P., Danchin, A., and Bertin, P. (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* **40**, 20-36.

Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD - Visual Molecular Dynamics. *J Mol Graphics* **14**, 33-38.

Johansson, J., Dagberg, B., Richet, E., and Uhlin, B. E. (1998) H-NS and StpA proteins stimulate expression of the maltose regulon in *Escherichia coli*. *J Bacteriol.* **180**, 6117-6125.

Johansson, J., Eriksson, S., Sonden, B., Wai, S. N., and Uhlin, B. E. (2001) Heteromeric interactions among nucleoid-associated bacterial proteins: localization of StpA-stabilizing regions in H-NS of *Escherichia coli*. *J Bacteriol* **183**, 2343-7.

Johansson, J. and Uhlin, B. E. (1999) Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of *Escherichia coli*. *Proc Natl Acad Sci USA* **96**, 10776-81.

Jordi, B. J., Fielder, A. E., Burns, C. M., Hinton, J. C., Dover, N., Ussery, D. W., and Higgins, C. F. (1997) DNA binding is not sufficient for H-NS-mediated repression of *proU* expression. *J Biol Chem* **272**, 12083-90.

Jordi, B. J. and Higgins, C. F. (2000) The downstream regulatory element of the *proU* operon of *Salmonella typhimurium* inhibits open complex formation by RNA polymerase at a distance. *J. Biol. Chem.* **275**, 12123-12128.

Jordi, B. J., Owen-Hughes, T. A., Hulton, C. S., and Higgins, C. F. (1995) DNA twist, flexibility and transcription of the osmoregulated *proU* promoter of *Salmonella typhimurium*. *Embo J* **14**, 5690-5700.

Kabsch, W. (1978) A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Cryst.* **A34**, 827-828.

Karem, K. and Foster, J. W. (1993) The influence of DNA topology on the environmental regulation of a pH-regulated locus in *Salmonella typhimurium*. *Mol Microbiol.* **10**, 75-86.

Kawula, T. H. and Orndorff, P. E. (1991) Rapid site-specific DNA inversion in *Escherichia coli* mutants lacking the histone-like protein H-NS. *J. Bacteriol.* **173**, 4116-4123.

Kay, L. E., Keifer P., and Saarinen T. (1992a) Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J Am Chem Soc* **114**, 10663-10665.

Kay, L. E. and Muhandiram, D. R. (1994) Gradient Enhanced Triple Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *J Magn Reson Series B* **103**, 203-216.

Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., and Torchia, D. A. (1992b) Pulse sequences for removal of the effects of cross-correlation between dipolar and chemi-

cal shift anisotropy relaxation mechanism on the measurement of heteronuclear T_1 and T_2 values in proteins. *J Magn Reson* **97**, 359-375.

Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979.

Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chains ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J Magn Reson Series B* **101**, 333-337.

Kostrewa, D., Granzin, J., Stock, D., Choe, H. W., Labahn, J., and Saenger, W. (1992) Crystal structure of the factor for inversion stimulation FIS at 2.0 A resolution. *J. Mol. Biol.* **226**, 209-226.

Kraulis, P. J. (1989) ANSIG: A Program for the Assignment of Protein ^{1}H 2D NMR spectra by Interactive Graphics. *J Magn Reson* **24**, 624-633.

Kraulis, P. J. (1994) Protein three-dimensional structure determination and sequence-specific assignment of ^{13}C and ^{15}N-separated NOE data. A novel real-space *ab initio* approach. *J. Mol. Biol.* **243**, 696-718.

Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) Measurement of HN-Hα *J* couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J Biomol. NMR* **4**, 871-878.

Landini, P. and Zehnder, A. J. (2002) The global regulatory *hns* gene negatively affects adhesion to solid surfaces by anaerobically grown *Escherichia coli* by modulating expression of flagellar genes and lipopolysaccharide production. *J Bacteriol.* **184**, 1522-1529.

Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol. NMR* **8**, 477-486.

Lease, R. A. and Belfort, M. (2000) Riboregulation by DsrA RNA: trans-actions for global economy. *Mol Microbiol.* **38**, 667-672.

Lease, R. A., Cusick, M. E., and Belfort, M. (1998) Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc Natl Acad Sci USA* **95**, 12456-61.

Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) Automated assignment of ambiguous nuclear Overhauser effects with ARIA. *Methods Enzymol.* **339**, 71-90.

Liu, Q. and Richardson, C. C. (1993) Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **90**, 1761-1765.

Lupas, A. (1997) Predicting coiled-coil regions in proteins. *Curr. Opin. Struct. Biol.* **7**, 388-393.

Madrid, C., Nieto, J. M., Paytubi, S., Falconi, M., Gualerzi, C. O., and Juarez, A. (2002) Temperature- and H-NS-dependent regulation of a plasmid-encoded virulence operon expressing *Escherichia coli* hemolysin. *J Bacteriol.* **184**, 5058-5066.

Mayer, O., Waldsich, C., Grossberger, R., and Schroeder, R. (2002) Folding of the *td* pre-RNA with the help of the RNA chaperone StpA. *Biochem. Soc Trans.* **30**, 1175-1180.

McGhee, J. D. and von Hippel, P. H. (1974) Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **86**, 469-489.

McGovern, V., Higgins, N. P., Chiz, R. S., and Jaworski, A. (1994) H-NS overexpression induces an artificial stationary phase by silencing global transcription. *Biochimie* **76**, 1019-1029.

More, M. I., Pohlman, R. F., and Winans, S. C. (1996) Genes encoding the pKM101 conjugal mating pore are negatively regulated by the plasmid-encoded KorA and KorB proteins. *J Bacteriol.* **178**, 4392-4399.

Nasser, W. and Reverchon, S. (2002) H-NS-dependent activation of pectate lyases synthesis in the phytopathogenic bacterium *Erwinia chrysanthemi* is mediated by the PecT repressor. *Mol Microbiol* **43**, 733-48.

Nieto, J. M., Madrid, C., Miquelay, E., Parra, J. L., Rodriguez, S., and Juarez, A. (2002) Evidence for direct protein-protein interaction between members of the enterobacterial Hha/YmoA and H-NS families of proteins. *J Bacteriol* **184**, 629-35.

Nieto, J. M., Madrid, C., Prenafeta, A., Miquelay, E., Balsalobre, C., Carrascal, M., and Juarez, A. (2000) Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. *Mol Gen Genet* **263**, 349-58.

Nilges, M. and O'Donoghue, S. I. (1998) Ambiguous NOEs and automated NOE assignment. *Prog. NMR Spect.* **32**, 107-139.

Nilges, M., Macias, M. J., O'Donoghue, S. I., and Oschkinat, H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β-spectrin. *J Mol. Biol.* **269**, 408-422.

Nye, M. B. and Taylor, R. K. (2003) *Vibrio cholerae* H-NS domain structure and function with respect to transcriptional repression of ToxR regulon genes reveals differences among H-NS family members. *Mol Microbiol.* **50**, 427-444.

O'Gara, J. P. and Dorman, C. J. (2000) Effects of local transcription and H-NS on inversion of the *fim* switch of *Escherichia coli*. *Mol. Microbiol.* **36**, 457-466.

Ohta, T., Ueguchi, C., and Mizuno, T. (1999) *rpoS* function is essential for *bgl* silencing caused by C-terminally truncated H-NS in *Escherichia coli*. *J Bacteriol* **181**, 6278-83.

Ohyama, T. (2001) Intrinsic DNA bends: an organizer of local chromatin structure for transcription. *Bioessays* **23**, 708-715.

Owen-Hughes, T. A., Pavitt, G. D., Santos, D. S., Sidebotham, J. M., Hulton, C. S., Hinton, J. C., and Higgins, C. F. (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* **71**, 255-265.

Palmer A.G., Cavanagh J., Wright P.E., and Rance M. (1991) Sensitivity Improvement in Proton-Detected Two-Dimensional Heteronuclear NMR Spectroscopy. *J Magn Reson* **93**, 151-170.

Polshakov, V. I., Frenkiel, T. A., Birdsall, B., Soteriou, A., and Feeney, J. (1995) Determination of stereospecific assignments, torsion-angle constraints and rotamer populations in proteins using the program AngleSearch. *J Magn Reson Series B* **108**, 31-43.

Poore, C. A. and Mobley, H. L. (2003) Differential regulation of the *Proteus mirabilis* urease gene cluster by UreR and H-NS. *Microbiology* **149**, 3383-3394.

Prosseda, G., Falconi, M., Giangrossi, M., Gualerzi, C. O., Micheli, G., and Colonna, B. (2004) The *virF* promoter in *Shigella*: more than just a curved DNA stretch. *Mol Microbiol.* **51**, 523-537.

Rajkumari, K. and Gowrishankar, J. (2001) In vivo expression from the RpoS-dependent P1 promoter of the osmotically regulated *proU* operon in *Escherichia coli* and *Salmonella enterica* serovar Typhimurium: activation by *rho* and *hns* mutations and by cold stress. *J. Bacteriol.* **183**, 6543-6550.

Renzoni, D., Esposito, D., Pfuhl, M., Hinton, J. C., Higgins, C. F., Driscoll, P. C., and Ladbury, J. E. (2001) Structural characterization of the N-terminal oligomerisation domain of the bacterial chromatin-structuring protein, H-NS. *J Mol Biol* **306**, 1127-37.

Reusch, R. N., Shabalin, O., Crumbaugh, A., Wagner, R., Schroder, O., and Wurm, R. (2002) Post-translational modification of *E. coli* histone-like protein H-NS and bovine histones by short-chain poly-(R)-3-hydroxybutyrate (cPHB). *FEBS Lett.* **527**, 319-322.

Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**, 1295-1306.

Rimsky, S., Zuber, F., Buckle, M., and Buc, H. (2001) A molecular mechanism for the repression of transcription by the H-NS protein. *Mol Microbiol.* **42**, 1311-1323.

Roberts, G. C. K. (1993) NMR of Macromolecules, A Practical Approach.

Robinow, C. F. (1956) The chromatin bodies of bacteria. *Bacteriol. Rev.* **20**, 207-242.

Schroder, O., Tippner, D., and Wagner, R. (2001) Toward the three-dimensional structure of the *Escherichia coli* DNA-binding protein H-NS: A CD and fluorescence study. *Biochem Biophys Res Commun* **282**, 219-27.

Schroder, O. and Wagner, R. (2002) The bacterial regulatory protein H-NS--a versatile modulator of nucleic acid structures. *Biol. Chem.* **383**, 945-960.

Sherratt, D. J. (2003) Bacterial chromosome dynamics. *Science* **301**, 780-785.

Shi, X. and Bennett, G. N. (1994) Plasmids bearing *hfq* and the *hns*-like gene *stpA* complement *hns* mutants in modulating arginine decarboxylase gene expression in *Escherichia coli*. *J Bacteriol* **176**, 6769-75.

Shiga, Y., Sekine, Y., Kano, Y., and Ohtsubo, E. (2001) Involvement of H-NS in transpositional recombination mediated by IS1. *J. Bacteriol.* **183**, 2476-2484.

Shimada, H., Wada, T., Handa, H., Ohta, H., Mizoguchi, H., Nishimura, K., Masuda, T., Shioi, Y., and Takamiya, K. (1996) A transcription factor with a leucine-zipper motif involved in light-dependent inhibition of expression of the *puf* operon in the photosynthetic bacterium *Rhodobacter sphaeroides*. *Plant Cell Physiol* **37**, 515-522.

Shindo, H., Iwaki, T., Ieda, R., Kurumizaka, H., Ueguchi, C., Mizuno, T., Morikawa, S., Nakamura, H., and Kuboniwa, H. (1995) Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett.* **360**, 125-131.

Shindo, H., Ohnuki, A., Ginba, H., Katoh, E., Ueguchi, C., Mizuno, T., and Yamazaki, T. (1999) Identification of the DNA binding surface of H-NS protein from *Escherichia coli* by heteronuclear NMR spectroscopy. *FEBS Lett.* **455**, 63-69.

Smyth, C. P. (1999) Oligomerisation of the Chromatin-Structuring Protein, H-NS. PhD thesis, University of London.

Smyth, C. P., Lundback, T., Renzoni, D., Siligardi, G., Beavil, R., Layton, M., Sidebotham, J. M., Hinton, J. C., Driscoll, P. C., Higgins, C. F., and Ladbury, J. E. (2000) Oligomerization of the chromatin-structuring protein H-NS. *Mol Microbiol* **36**, 962-72.

Sonden, B. and Uhlin, B. E. (1996) Coordinated and differential expression of histone-like proteins in *Escherichia coli*: regulation and function of the H-NS analog StpA. *EMBO J.* **15**, 4970-4980.

Sonnenfield, J. M., Burns, C. M., Higgins, C. F., and Hinton, J. C. (2001) The nucleoid-associated protein StpA binds curved DNA, has a greater DNA-binding affinity than H-NS and is present in significant levels in *hns* mutants. *Biochimie* **83**, 243-9.

Soutourina, O. A., Krin, E., Laurent-Winter, C., Hommais, F., Danchin, A., and Bertin, P. N. (2002) Regulation of bacterial motility in response to low pH in *Escherichia coli*: the role of H-NS protein. *Microbiology* **148**, 1543-51.

Spera, S. and Bax, A. (1991) Empirical Correlation between Protein Backbone Conformation and Cα and Cβ ^{13}C Nuclear Magnetic Resonance Chemical Shifts. *J Am Chem Soc* **113**, 5490-5492.

Sperandio, V., Mellies, J. L., Delahay, R. M., Frankel, G., Crawford, J. A., Nguyen, W., and Kaper, J. B. (2000) Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler. *Mol. Microbiol.* **38**, 781-793.

Spurio, R., Durrenberger, M., Falconi, M., La Teana, A., Pon, C. L., and Gualerzi, C. O. (1992) Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. *Mol. Gen. Genet.* **231**, 201-211.

Spurio, R., Falconi, M., Brandi, A., Pon, C. L., and Gualerzi, C. O. (1997) The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *Embo J* **16**, 1795-1805.

Starcic-Erjavec, M., van Putten, J. P., Gaastra, W., Jordi, B. J., Grabnar, M., and Zgur-Bertok, D. (2003) H-NS and Lrp serve as positive modulators of *traJ* expression from the *Escherichia coli* plasmid pRK100. *Mol Genet. Genomics* **270**, 94-102.

Stevens, F. J. (1986) Analysis of protein-protein interaction by simulation of small-zone size-exclusion chromatography: application to an antibody-antigen association. *Biochemistry* **25**, 981-993.

Stevens, F. J. (1989) Analysis of protein-protein interaction by simulation of small-zone size exclusion chromatography. Stochastic formulation of kinetic rate contributions to observed high-performance liquid chromatography elution characteristics. *Biophys. J.* **55**, 1155-1167.

Sturtevant, J. M. (1977) Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. U. S. A* **74**, 2236-2240.

Szyperski, T., Luginbuhl, P., Otting, G., Guntert, P., and Wuthrich, K. (1993) Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J Biomol. NMR* **3**, 151-164.

Tendeng, C., Badaut, C., Krin, E., Gounon, P., Ngo, S., Danchin, A., Rimsky, S., and Bertin, P. (2000) Isolation and characterization of *vicH*, encoding a new pleiotropic regulator in *Vibrio cholerae*. *J. Bacteriol.* **182**, 2026-2032.

Tendeng, C. and Bertin, P. N. (2003) H-NS in Gram-negative bacteria: a family of multifaceted proteins. *Trends Microbiol.* **11**, 511-518.

Tendeng, C., Krin, E., Soutourina, O. A., Marin, A., Danchin, A., and Bertin, P. N. (2003) A Novel H-NS-like protein from an antarctic psychrophilic bacterium reveals a crucial role for the N-terminal domain in thermal stability. *J. Biol. Chem.* **278**, 18754-18760.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Tippner, D. and Wagner, R. (1995) Fluorescence analysis of the *Escherichia coli* transcription regulator H-NS reveals two distinguishable complexes dependent on binding to specific or nonspecific DNA sites. *J Biol Chem* **270**, 22243-7.

Tupper, A. E., Owen-Hughes, T. A., Ussery, D. W., Santos, D. S., Ferguson, D. J., Sidebotham, J. M., Hinton, J. C., and Higgins, C. F. (1994) The chromatin-associated protein H-NS alters DNA topology *in vitro*. *Embo J* **13**, 258-268.

Ueguchi, C., Kakeda, M., and Mizuno, T. (1993) Autoregulatory expression of the *Escherichia coli hns* gene encoding a nucleoid protein: H-NS functions as a repressor of its own transcription. *Mol Gen. Genet.* **236**, 171-178.

Ueguchi, C. and Mizuno, T. (1993) The *Escherichia coli* nucleoid protein H-NS functions directly as a transcriptional repressor. *Embo J* **12**, 1039-1046.

Ueguchi, C., Seto, C., Suzuki, T., and Mizuno, T. (1997) Clarification of the dimerization domain and its functional significance for the *Escherichia coli* nucleoid protein H-NS. *J Mol Biol* **274**, 145-51.

Ueguchi, C., Suzuki, T., Yoshida, T., Tanaka, K., and Mizuno, T. (1996) Systematic mutational analysis revealing the functional domain organization of *Escherichia coli* nucleoid protein H-NS. *J Mol Biol* **263**, 149-62.

Umanski, T., Rosenshine, I., and Friedberg, D. (2002) Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*. *Microbiology* **148**, 2735-2744.

Ussery, D., Larsen, T. S., Wilkes, K. T., Friis, C., Worning, P., Krogh, A., and Brunak, S. (2001) Genome organisation and chromatin structure in *Escherichia coli*. *Biochimie* **83**, 201-12.

Ussery, D. W., Hinton, J. C., Jordi, B. J., Granum, P. E., Seirafi, A., Stephen, R. J., Tupper, A. E., Berridge, G., Sidebotham, J. M., and Higgins, C. F. (1994) The chromatin-associated protein H-NS. *Biochimie* **76**, 968-980.

Vuister, G. W. and Bax, A. (1992) Resolution Enhancement and Spectral Editing of Uniformly ^{13}C Enriched Proteins by Homonuclear Broadband ^{13}C Decoupling. *J Magn Reson* **98**, 428-435.

Vuister, G. W. and Bax, A. (1993) Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond $J_{HNH\alpha}$ Coupling Constants in ^{15}N-Enriched Proteins. *J Am Chem Soc* **115**, 7772-7777.

Waldsich, C., Grossberger, R., and Schroeder, R. (2002) RNA chaperone StpA loosens interactions of the tertiary structure in the *td* group I intron *in vivo*. *Genes Dev.* **16**, 2300-2312.

White, S. W., Wilson, K. S., Appelt, K., and Tanaka, I. (1999) The high-resolution structure of DNA-binding protein HU from *Bacillus stearothermophilus*. *Acta Crystallogr.* D. Biol. *Crystallogr.* **55 (Pt 4)**, 801-809.

Williams, R. M. and Rimsky, S. (1997) Molecular aspects of the *E. coli* nucleoid protein, H-NS: a central controller of gene regulatory networks. *FEMS Microbiol Lett* **156**, 175-85.

Williams, R. M., Rimsky, S., and Buc, H. (1996) Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *J Bacteriol* **178**, 4335-43.

Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131-137.

Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol. Biol.* **222**, 311-333.

Woody, R. W. (1995) Circular dichroism. *Methods Enzymol.* **246**, 34-71.

Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids.

Wuthrich, K. (2001) The way to NMR structures of proteins. *Nat. Struct. Biol.* **8**, 923-925.

Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) 2-Dimensional NMR experiments for correlating $^{13}C\beta$ and $^{1}H\delta/^{1}H\varepsilon$ chemical-shifts of aromatic residues in $^{13}C\beta$-labeled proteins via scalar couplings. *J Am Chem Soc* **115**, 11054-11055.

Yamazaki, T., Lee, W., Revington, M., Mattiello, D., Dahlquist, F. W., Arrowsmith, C. H., and Kay, L. E. (1994) An HNCA pulse scheme for the backbone assignment of ^{15}N, ^{13}C, ^{2}H-labeled proteins: application to a 37-kDa Trp repressor-DNA complex. *J Am Chem Soc* **116**, 6464-6465.

Yee, A., Chang, X., Pineda-Lucena, A., Wu, B., Semesi, A., Le, B., Ramelot, T., Lee, G. M., Bhattacharyya, S., Gutierrez, P., Denisov, A., Lee, C. H., Cort, J. R., Kozlov,

G., Liao, J., Finak, G., Chen, L., Wishart, D., Lee, W., McIntosh, L. P., Gehring, K., Kennedy, M. A., Edwards, A. M., and Arrowsmith, C. H. (2002) An NMR approach to structural proteomics. *Proc. Natl. Acad. Sci. U. S. A* **99**, 1825-1830.

Zhang, A. and Belfort, M. (1992) Nucleotide sequence of a newly-identified *Escherichia coli* gene, *stpA*, encoding an H-NS-like protein. *Nucleic Acids Res.* **20**, 6735-

Zhang, A., Derbyshire, V., Salvo, J. L., and Belfort, M. (1995) *Escherichia coli* protein StpA stimulates self-splicing by promoting RNA assembly *in vitro*. *RNA*. **1**, 783-793.

Zhang, A., Rimsky, S., Reaban, M. E., Buc, H., and Belfort, M. (1996) *Escherichia coli* protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *Embo J* **15**, 1340-9.

Zuber, F., Kotlarz, D., Rimsky, S., and Buc, H. (1994) Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol Microbiol.* **12**, 231-240.